

CONSISTENT PARTIAL LEAST SQUARES PATH MODELING¹

Theo K. Dijkstra

Faculty of Economics and Business, University of Groningen, Nettelbosje 2,
9747 AE Groningen THE NETHERLANDS {t.k.dijkstra@rug.nl}

Jörg Henseler

Faculty of Engineering Technology, University of Twente, Drienerlolaan 5,
7522 NB Enschede THE NETHERLANDS {j.henseler@utwente.nl}
NOVA IMS, Universidade Nova de Lisboa, 1070-312 Lisbon PORTUGAL {jhenseler@isegi.unl.pt}

This paper resumes the discussion in information systems research on the use of partial least squares (PLS) path modeling and shows that the inconsistency of PLS path coefficient estimates in the case of reflective measurement can have adverse consequences for hypothesis testing. To remedy this, the study introduces a vital extension of PLS: consistent PLS (PLSc). PLSc provides a correction for estimates when PLS is applied to reflective constructs: The path coefficients, inter-construct correlations, and indicator loadings become consistent. The outcome of a Monte Carlo simulation reveals that the bias of PLSc parameter estimates is comparable to that of covariance-based structural equation modeling. Moreover, the outcome shows that PLSc has advantages when using non-normally distributed data. We discuss the implications for IS research and provide guidelines for choosing among structural equation modeling techniques.

Keywords: PLS, consistent partial least squares, SEM, variance-based structural equation modeling, Monte Carlo simulation

Introduction

There has been a proliferation of structural equation modeling (SEM) in information systems and other social science and business disciplines (Gefen et al. 2011). Researchers have embraced its ability to model latent variables, correct for measurement errors, and estimate parameters of entire theo-

ries simultaneously. Two families of structural equation modeling techniques prevail (Chin 1998a): covariance-based SEM and variance-based SEM. Variance-based SEM involves many different techniques, such as regression on sum scores or principal components (Tenenhaus 2008), generalized structured component analysis (Henseler 2012; Hwang and Takane 2004), and partial least squares path modeling (PLS; Wold 1982). Among variance-based SEM techniques, PLS has been regarded as the “most fully developed and general system” (McDonald 1996, p. 240) and has been called a “silver bullet” (Hair et al. 2011). IS research has relied heavily on PLS as a method of statistical analysis (Marcoulides and Saunders 2006), and *MIS Quarterly*, in particular, has published a large number of PLS applications (Ringle et al. 2012).

The use of variance-based SEM is not without disadvantages. All variance-based structural equation modeling techniques—

¹Ron Thompson was the accepting senior editor for this paper. Christian Ringle served as the associate editor.

The appendices for this paper are located in the “Online Supplements” section of the *MIS Quarterly*’s website (<http://www.misq.org>).

© 2015 The Authors. Published by the Management Information Systems Research Center at the University of Minnesota. This is an open access article under the terms of the Creative Commons Attribution CC BY License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited.

PLS being no exception—approximate common factors using linear composites of observed variables. Accordingly, the obtained construct scores cannot be completely free from measurement error (for an excellent explanation of why PLS only partially corrects for random measurement error, see Rigdon 2012), and inter-construct correlations are subject to attenuation (Goodhue et al. 2012b; McDonald 1996). Wold (1982) noted that PLS estimates—in particular, path coefficients and indicator loadings—are not consistent, but only “consistent at large”² (i.e., both the number of observations and of reflective indicators must be large). Although researchers have been aware of PLS’s lack of consistency as an estimator of common factor models for over three decades (Fornell and Bookstein 1982), the discussion of its consequences has only recently gained momentum (Goodhue et al. 2012a, 2012b; Henseler, Ringle, and Sarstedt 2014; Marcoulides et al. 2012; Ringle et al. 2012). Researchers have acknowledged the lack of consistency in the sense that

Parameter estimates for paths between observed variables and latent variable proxies are biased upward in PLS (away from zero), while parameter estimates for paths between proxies are attenuated (Gefen et al. 2011, p. vi).

Such an underestimation of effects has the potential to negatively influence a method’s ability to detect a significant effect if it indeed exists in the population (the statistical power). For fields of research relying heavily on PLS and other variance-based SEM techniques as a means of analyzing relationships between reflective constructs, a potential lack of statistical power would be a concern. Since PLS analyses “often serve as a basis for theory development, promising research avenues might have been overlooked” (Hair, Sarstedt et al. 2012, p. 424). As a consequence, methodological papers exploring the characteristics of PLS, such as Goodhue et al. (2006, 2012b), Lu et al. (2011), or Reinartz et al. (2009), have often focused on the statistical power thereof. Monte Carlo simulations of the aforementioned papers largely confirmed that the power of PLS is comparable to competing techniques, such as covariance-based SEM or regression on sum scores, which suggests that researchers need not worry about the statistical power of PLS. However, none of the extant simulation studies have investigated whether substantial collinearity in the structural model influences the adequacy and statistical power of path coefficient estimates obtained through variance-based SEM.

²Wold noted that the PLS estimators for path coefficients and loadings do not tend in probability to the true values when the sample size increases, so they are not consistent in the traditional sense. He pointed out, however, that they will tend to the true values when the number of reflective indicators increases as well. He introduced the term *consistency at large* to capture both conditions.

Thus far, rather little attention has been paid to the possibility that PLS, if used to estimate common factor models, overestimates path coefficients. Inflated path coefficients increase the danger of Type I errors (false positives), which means that an effect may be considered significant even though it does not actually exist in the population. Monte Carlo simulations either do not investigate the false positives (e.g., Lu et al. 2011; Reinartz et al. 2009) or come to the conclusion that PLS does not produce more false positives for which the chosen significance level allows (Goodhue et al. 2006, 2012b). Only a few papers have reported that variance-based SEM can find effects that are not actually present. Rigdon (1994, p. 379) explicitly warned that “random measurement error, if ignored, can have a profound impact on modeling results.” Henseler (2012) showed that variance-based SEM can inflate path coefficients in the case of mediation, and Goodhue et al. (2011) demonstrated that regression on sum scores and PLS provide exaggerated t-values if multicollinearity is present among independent latent variables.

In social and business science, false positives are often regarded as a more severe problem than false negatives (Cohen 1988). This opinion is also the reason why confidence levels of 95 percent or higher are often required, whereas a statistical power level of 80 percent is deemed acceptable (Goodhue et al. 2012b). Therefore, cases in which variance-based SEM detects a nonzero path coefficient, even though the effect is actually zero, deserve further analysis. We demonstrate the problem for a simple structural model with one endogenous latent variable and two correlated exogenous latent variables, as depicted in Figure 1. All latent variables have reflective measurement models whose indicators are measured with random error. The equation of the structural model is $\eta = \beta_1 \cdot \xi_1 + \beta_2 \cdot \xi_2 + \delta$. It can be algebraically derived (see Appendix A) that, if the true path coefficient β_1 is zero, the estimate $\hat{\beta}_1$ of the direct effect of the latent variable ξ_1 on the endogenous latent variable η is not necessarily zero, even though it is determined based on population data. In this case, the path coefficient $\hat{\beta}_1$ of variance-based SEM can be determined as follows (tildes denote the construct scores):

$$\hat{\beta}_1 = \frac{\overbrace{\text{correlation}(\tilde{\xi}_1, \tilde{\xi}_2)}^{\text{degree of multicollinearity}}}{1 - \text{correlation}^2(\tilde{\xi}_1, \tilde{\xi}_2)} \cdot \text{correlation}(\tilde{\xi}_2, \tilde{\eta}) \cdot \frac{1 - \text{reliability}(\tilde{\xi}_2)}{\text{reliability}(\tilde{\xi}_2)} \quad (1)$$

Although the true parameter value β_1 is zero, the estimate $\hat{\beta}_1$ does not necessarily equate to zero. Its value depends on the degree of collinearity (as quantified by the correlation between the exogenous variables’ scores, $\tilde{\xi}_1$ and $\tilde{\xi}_2$), the corre-

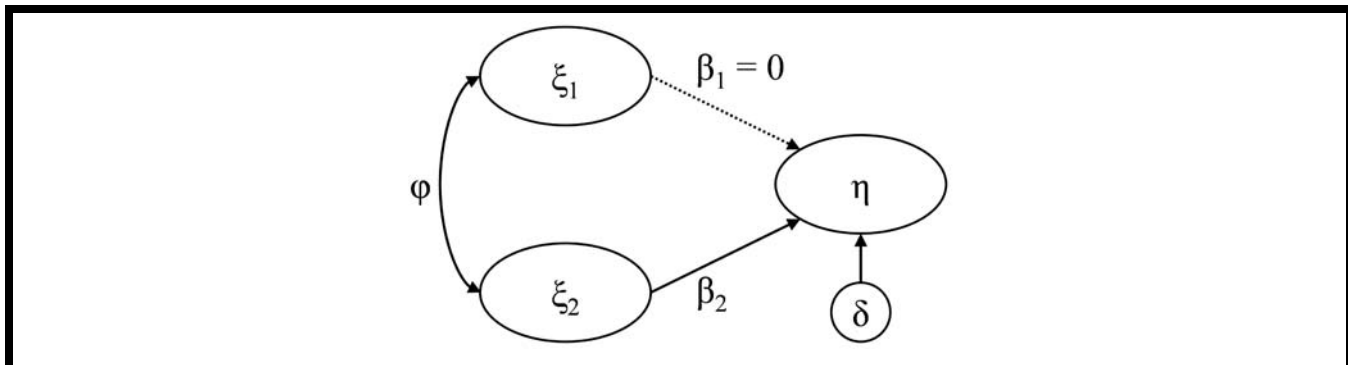


Figure 1. A Model that Illustrates the Inconsistency Problem of Variance-Based SEM

lation between the scores of the second exogenous variable ξ_2 and the endogenous variable η , and the reliability of the second exogenous variable's scores ($\hat{\xi}_2$). As can be derived from Equation 1, the bias is larger the less reliable the second exogenous variable's scores, the larger the correlation between the scores of the second exogenous variable ξ_2 and the endogenous variable η , and the larger the collinearity. The effect can only be correctly estimated as zero if ξ_2 is measured without any measurement error (which implies that its reliability equals one), if the second exogenous variable ξ_2 and the endogenous variable η are completely uncorrelated, or if there is no multicollinearity at all.

When using variance-based SEM, misleading estimates for structural equation models with reflective measurement models are neither innocuous nor infrequent (Goodhue et al. 2011). If variance-based SEM underestimates the true parameter, Type II errors are likely; if variance-based SEM overestimates the true parameter, the Type I error is inflated. Inconsistent estimates can lead to erroneous conclusions and must be considered potentially harmful to hypothesis testing. The often-held belief that an inconsistency of estimates is unproblematic as long as researchers are only interested in the sheer existence of an effect but not its size is wrong. Consistent estimates must thus be considered the *sine qua non* of hypothesis testing.

In light of PLS's lack of consistency when estimating common factor models, researchers can choose between continuing to use PLS while acknowledging its deficiencies (Marcoulides et al. 2012), avoiding PLS completely (Rönkkö and Ylitalo 2010), or correcting its estimates in some way (Dijkstra 1981; Goodhue et al. 2012b). Given that PLS has some undisputed advantages, such as its ability to estimate composite models (Henseler et al. 2014), its predictive capabilities (Becker, Rai et al. 2013) or its lack of convergence problems (Henseler 2010), it appears worthwhile to

pursue the third option. The present article thus introduces a pivotal improvement to traditional PLS, namely, consistent partial least squares (PLSc). PLSc corrects estimates of reflectively measured constructs using a novel reliability coefficient ρ_A . It overcomes traditional PLS' consistency problems when estimating common factor models in the sense that it consistently estimates the path coefficients, inter-construct correlations, and indicator loadings.

We conducted a Monte Carlo simulation to explore PLSc's performance. The outcome of this computational experiment revealed that the bias of PLSc is comparable with covariance-based SEM, and that PLSc is nearly as efficient as FIML (full-information maximum likelihood) for medium to large sample sizes. Similar to covariance-based SEM, PLSc avoids the excessive amount of Type I and Type II errors that can occur if traditional PLS or regression on sum scores is applied to estimate structural equation models with reflective measurement models. Moreover, we find that the statistical power of PLSc is only slightly lower than that of FIML, and that PLSc does not produce an undue number of false positives when the data are not normally distributed. We discuss the implications of our findings for IS research and provide guidelines for choosing between structural equation modeling techniques.

Consistent Partial Least Squares Path Modeling

A Consistent Reliability Coefficient for PLS

Consistent PLS (PLSc) was first mathematically developed by Dijkstra (1981, 2010, 2011, 2014). The logic of PLSc departs from the suggestion of Goodhue et al. (2012b, p. 996) "to 'adjust' the attenuated regression path estimates to the correct

value using the reliability of the constructs” in that it uses Nunnally and Bernstein’s equation (1994, pp. 241, 257):

$$r_{xy}^c = \frac{r_{xy}}{\sqrt{r_{xx} \cdot r_{yy}}} \quad (2)$$

in which r_{xy}^c is the consistent correlation between two constructs x and y ; r_{xy} is the correlation between the constructs’ scores; and r_{xx} and r_{yy} represent the constructs’ reliability. Equation 2 underlines the importance of determining the reliability of PLS’s construct scores: Consistent estimates for correlations between reflectively measured latent variables as well as derived coefficients, such as partial correlations, part correlations, or path coefficients, can only be obtained if viable reliability coefficients are available. The major challenge of this approach is to consistently estimate the constructs’ reliability. As we will show, the two reliability coefficients that are typically applied in conjunction with PLS, namely Cronbach’s coefficient alpha (Cronbach 1951) and composite reliability (Chin 2010), are not consistent themselves. Thus, even if researchers corrected for attenuation, the lack of a consistent reliability coefficient would prevent them from achieving their aim (i.e., obtaining consistent estimates of the relationships between latent variables).

Cronbach’s alpha evaluates the indicator variances and covariances to determine the internal consistency reliability. It is a consistent estimate of a linear composite’s reliability if the composite’s indicators form a unidimensional set (Gerbing and Anderson 1988) and their covariances are equal (tau-equivalence; see Lord and Novick 1968). If this assumption of tau-equivalence is not met or if the sample size is small, Cronbach’s alpha will underestimate the reliability (Sijtsma 2009; Yuan and Bentler 2002). These characteristics of Cronbach’s alpha interfere with PLS and other variance-based SEM techniques in two ways. First, variance-based SEM is preferred to covariance-based structural equation modeling when the sample size is small (Chin and Newsted 1999). Second, variance-based structural equation modeling neither tests nor ensures the tau-equivalence of indicators. Consequently, Cronbach’s alpha is unlikely to consistently estimate the reliability of PLS construct scores.

In reaction to Cronbach’s alpha’s inappropriateness, Chin (1998b) suggested the composite reliability ρ_c (Heise and Bohrnstedt 1970) as a more appropriate measure of reliability. To determine the reliability of construct scores, composite reliability assesses the indicator loadings. According to Chin (1998a), “ ρ_c is a closer approximation [of reliability] under the assumption that the parameter estimates are accurate.” However, the assumption of parameter accuracy is unlikely to hold, because the indicator loadings provided by PLS are

known to be upward-biased (Dijkstra 1983). Consequently, composite reliability is likely to overestimate the actual reliability of construct scores.

In light of the deficiencies of extant reliability coefficients, we propose a new and consistent reliability coefficient for PLS: ρ_A . Index “A” was chosen because the construct scores are generated by means of Mode A of PLS. The new reliability coefficient ρ_A resides on two pillars. First, unlike composite reliability, ρ_A evaluates a construct’s weights, not its loadings. This design choice was motivated by Dijkstra’s (2010) finding that the probability limits (roughly, the values obtained for the population) of the construct weights obtained by means of Mode A are proportional to the true loadings. Second, ρ_A is determined such that the off-diagonal elements of a latent variable’s indicator correlation matrix are reproduced as well as possible in a least squares sense. We define ρ_A as follows:

$$\rho_A := (\hat{\mathbf{w}}' \hat{\mathbf{w}})^2 \cdot \frac{\hat{\mathbf{w}}' (\mathbf{S} - \text{diag}(\mathbf{S})) \hat{\mathbf{w}}}{\hat{\mathbf{w}}' (\hat{\mathbf{w}} \hat{\mathbf{w}}' - \text{diag}(\hat{\mathbf{w}} \hat{\mathbf{w}}')) \hat{\mathbf{w}}} \quad (3)$$

In Equation 3, $\hat{\mathbf{w}}$ is the estimated weight vector of the latent variable (the dimension of $\hat{\mathbf{w}}$ is the number of indicators directly associated with the latent variable), and \mathbf{S} is the empirical covariance matrix of the latent variable’s indicators. Appendix B contains a detailed derivation of Equation 3. Having a consistent reliability measure for PLS construct scores available paves the way for PLSc.

The Four Steps of Consistent PLS

We introduce PLSc as a procedure consisting of four steps. (1) Traditional PLS is applied to provide latent variable scores and to estimate latent variable correlations and weights. (2) The new reliability coefficient ρ_A is determined for each reflective construct. (3) ρ_A can be used to correct the original latent variable correlations for attenuation and thus obtain consistent latent variable correlations. (4) Finally, consistent path coefficients are estimated in a least squares manner based on the consistent latent variable correlations. Figure 2 visualizes these constituting steps of PLSc. We describe each step in more detail below.

Step 1: Traditional PLS. As an input, PLSc requires correlations between latent variable scores and indicator weights, which are provided by traditional PLS. Latent variable scores (a side product of traditional PLS) are not essential for PLSc, but may be useful for predictive purposes (Becker et al. 2013a) or for add-on analyses, such as detecting unobserved

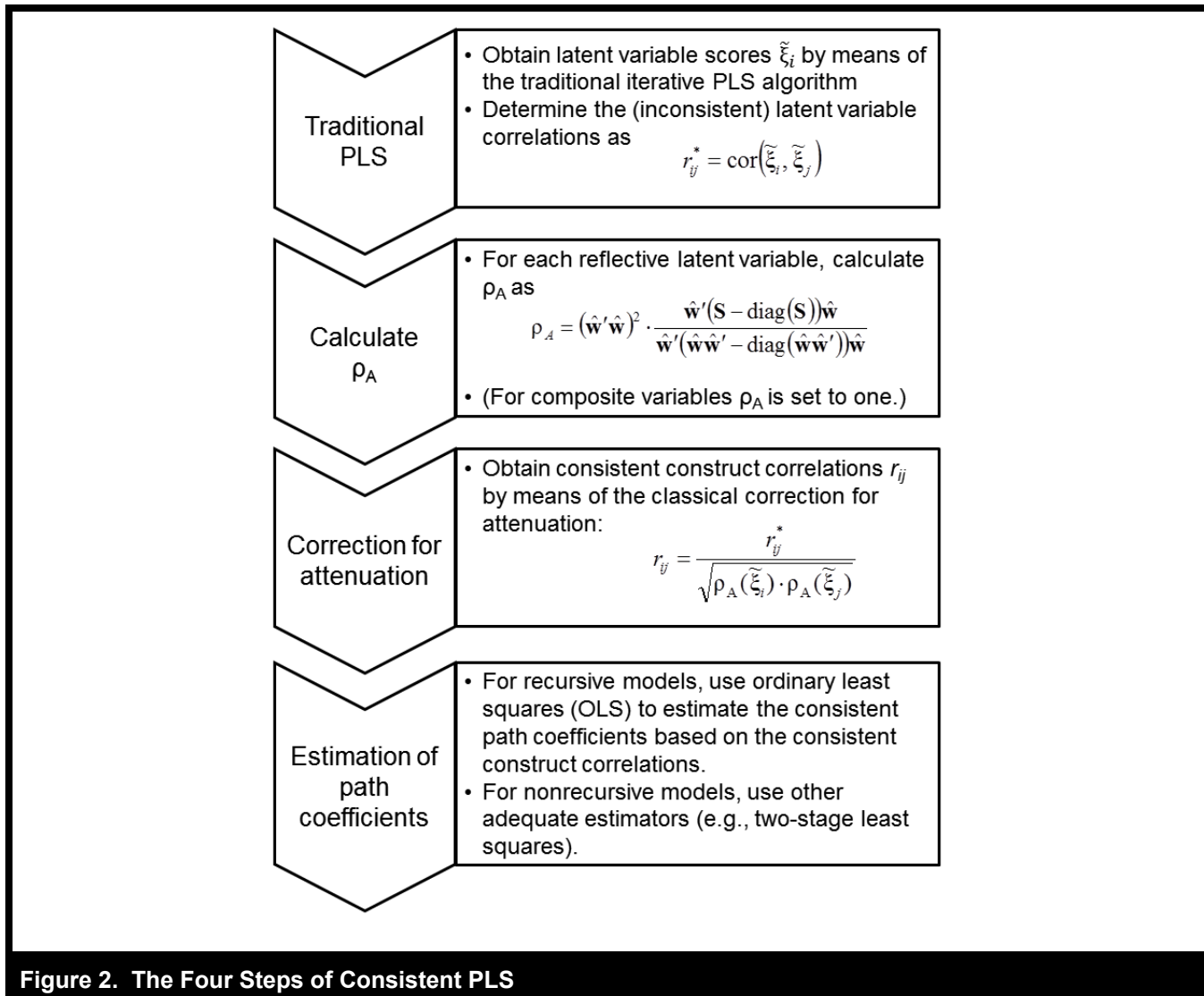


Figure 2. The Four Steps of Consistent PLS

heterogeneity (e.g., Becker, Rai, and Rigdon 2013). An integral part of the traditional iterative algorithm of PLS is the *inner approximation*, in which inner proxies are created for each latent variable (Lohmöller 1989; Wold 1982). This is a unique feature of PLS that does not emerge in other variance-based SEM techniques. The inner proxy of a focal latent variable is a certain weighted sum of the scores of other latent variables in the model. It forms the basis for estimating the indicator weights, which in turn are the major ingredient of ρ_A . The type of weighting depends on the inner weighting scheme. Three inner weighting schemes are typically employed in PLS: the centroid scheme, the factor scheme, and the path-weighting scheme. Whereas the factor scheme uses the correlations between construct scores as weights, the centroid scheme employs the sign of the correlations. The

path-weighting scheme uses regression weights for predictors and correlations for other constructs. It should be noted that the choice of weighting scheme does not affect the consistency and asymptotic normality of PLSc.

Step 2: Estimating reliability. For each reflective construct, the reliability of the construct scores must be estimated using the new reliability coefficient ρ_A , as expressed in Equation 3. Not only are the obtained reliability coefficients an integral part of PLSc, they should also be reported to the reviewers and readers of an empirical study to allow for comparisons with established thresholds of reliability (Nunnally and Bernstein 1994). It should be noted that it does not make sense to calculate ρ_A for constructs that are measured using formative measurement models or composite models, be-

cause, in these cases, the inter-indicator correlations are not informative regarding indicator reliability and thus random measurement errors. In the tradition of assuming that random measurement error³ does not affect formative measures and composites, ρ_A could be set equal to 1 when PLS is used to estimate composite models.

Step 3: Correction for attenuation. Consistent latent variable correlations are obtained using ρ_A to correct the original latent variable correlations for attenuation. For every pair of latent variable scores $\tilde{\eta}_i$ and $\tilde{\eta}_j$, the consistent correlation $\text{cor}(\eta_i, \eta_j)$ is calculated as follows:

$$\text{cor}(\eta_i, \eta_j) = \frac{\text{cor}(\tilde{\eta}_i, \tilde{\eta}_j)}{\sqrt{\rho_A(\tilde{\eta}_i) \cdot \rho_A(\tilde{\eta}_j)}} \quad (4)$$

The correction for attenuation is also required if one of the latent variables, η_i or η_j , is formative. If both latent variables are formative, no correction is required.

Step 4: Estimating consistent coefficients. Finally, consistent path coefficients are estimated based on the consistent latent variable correlations. If we make the common assumption that the path model is recursive, we can estimate the path coefficients for each endogenous latent variable by means of ordinary least squares (OLS) regression,⁴ that is, by using

$$\beta = \mathbf{R}_X^{-1} \mathbf{r}_{Xy} \quad (6)$$

In this equation, β denotes a vector of path coefficients, \mathbf{R}_X stands for the correlation matrix of the independent variables of the structural equation, and \mathbf{r}_{Xy} is the vector of correlations between the dependent variable and the independent variables. Equation 6 is the regular OLS equation for standardized coefficients based on correlations.

Consistent loadings. If the ρ_A and weight estimates are available, it is possible to estimate the loadings consistently (see Appendix B for the derivation). For the vector of loadings λ of a latent variable, we have:

$$\lambda = \hat{\mathbf{w}} \cdot \frac{\sqrt{\rho_A}}{\hat{\mathbf{w}}' \hat{\mathbf{w}}} \quad (5)$$

A Small Example Demonstrating the Viability of Consistent PLS

We applied a small example model, as depicted in Figure 3, to demonstrate the behavior of PLSc. The true model consisted of two moderately correlated latent variables with three indicators each and different loading patterns. We generated data that meets the true model's parameterization exactly.

We then estimated the model parameters both with traditional PLS and consistent PLS. The typical pattern of results was found for traditional PLS (see Dijkstra 1983): The loadings were overestimated, whereas the inter-construct correlation was attenuated. We could confirm the anticipated results for PLSc: It perfectly recovers the true parameter values; that is, the estimates for the loadings as well as the inter-construct correlation are equal to the true parameters. This example provides first evidence that PLSc is Fisher consistent, which means that the method provides the true parameters if applied to the population data.

This example also visualizes the differences between the reliability coefficients. Whereas Cronbach's alpha underestimates the reliability of both latent variable scores, composite reliability overestimates their reliability. By contrast, the new reliability coefficient ρ_A reproduces the actual reliability of construct scores exactly.

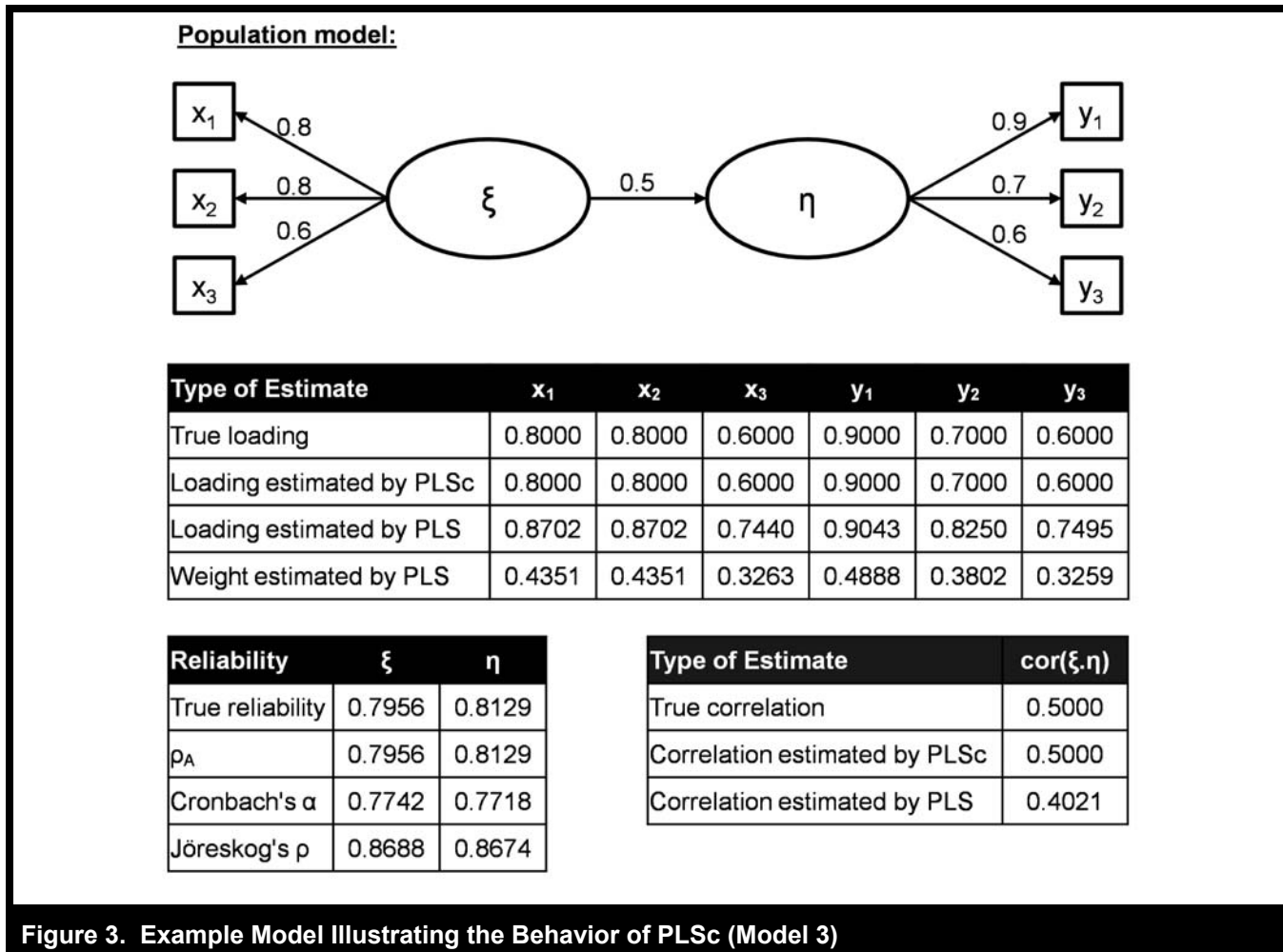
A Simulation Study

Design of the Monte Carlo Simulation

We conducted a Monte Carlo simulation to compare PLSc to plausible alternatives in the literature. MIS research typically relies on one of three techniques to test multiple hypotheses between latent variables: PLS, (multiple) regression on sum scores, and covariance-based SEM. Covariance-based SEM itself is a family of methods. Of these, we considered five techniques that we deemed to represent the spectrum of available techniques: full-information maximum likelihood (FIML), generalized least squares (GLS), weighted least squares (WLS), diagonally weighted least squares (DWLS), and unweighted least squares (ULS). We explored how—compared to the alternative techniques—PLSc performs in

³We acknowledge that random measurement error is also likely to be an issue in formative measurement, but the literature on formative measurement does not present possible solutions.

⁴In the case of non-recursive models, researchers should apply an appropriate simultaneous equations technique, of which two-stage least squares (2SLS) is the simplest choice, instead of OLS, to obtain consistent estimates (Dijkstra and Henseler 2015).



terms of convergence behavior, parameter consistency, and inference statistics. Convergence behavior reflects the certainty with which researchers can expect a particular technique to provide estimates: Does a technique lead to a solution? Parameter consistency refers to the bias of path coefficients: To what extent can the various techniques recover the true parameters? What parameter value can be expected if several samples are drawn from a population? Finally, inference statistics encompass the occurrence of Type I and Type II errors: Does a particular technique maintain the predefined alpha level? And what is its statistical power? We designed our computational experiment such that it could provide answers to all of these questions.

We selected a population model for the simulation, as depicted in Figure 4. Typically, a population model used in a Monte Carlo simulation should resemble those commonly estimated in applied research (Paxton et al. 2001). Our population model consists of five latent variables, which is

close to what Shah and Goldstein (2006) report as the average number of constructs used in covariance-based SEM. Moreover, it is a common configuration for simulation studies on structural equation modeling (see Goodhue et al. 2006; Jarvis et al. 2003). The reflective measurement models of our five constructs varied with respect to the number of indicators (2, 4, and 6), as well as construct reliability (approximately 0.7, 0.8, and 0.9). The structural model contained six effects, whose strengths were deliberately chosen:

1. The effect of ξ_2 on η_3 (γ_{23}) was zero and was not affected by multicollinearity. We expected all techniques to estimate this effect consistently.
2. The effect of η_3 on η_5 was medium-sized ($\beta_{35} = 0.35$) and not affected by multicollinearity. We expected PLSc and covariance-based SEM to estimate this effect consistently, whereas the estimates of traditional PLS and regression were expected to be attenuated. We assumed

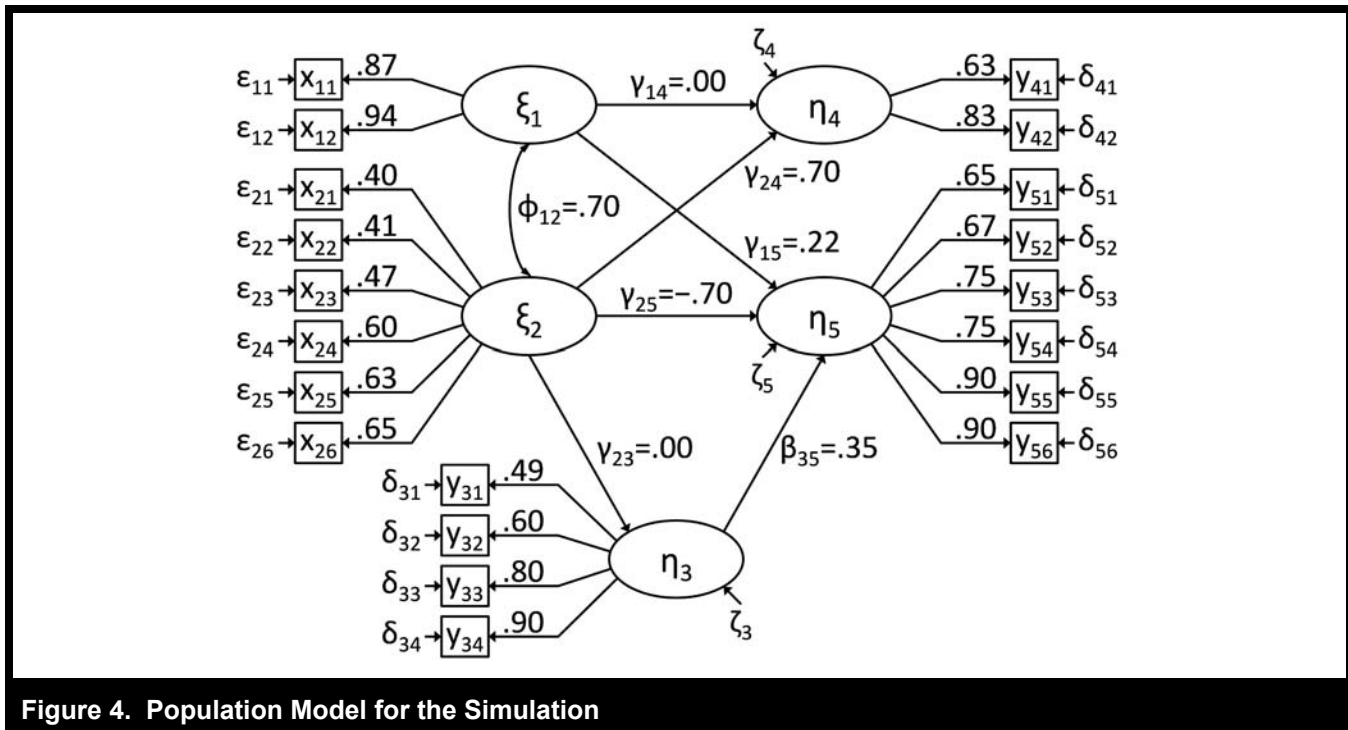


Figure 4. Population Model for the Simulation

- that this attenuation would be rather slight because the latent variables η_3 and η_5 are measured relatively reliably.
- The effect of ξ_2 on η_4 was strong ($\gamma_{24} = 0.70$) and substantially affected by multicollinearity due to the large correlation between ξ_1 and ξ_2 ($\phi_{12} = 0.70$). We expected PLSc and covariance-based SEM to estimate this effect consistently, whereas the estimates of traditional PLS and regression were expected to be attenuated. We assumed that this attenuation would be strong because the reliability of the latent variables ξ_2 and η_4 is only just acceptable.
 - The effect of ξ_2 on η_5 was strong but negative ($\gamma_{24} = -0.70$) and substantially affected by multicollinearity. We expected PLSc and covariance-based SEM to estimate this effect consistently, whereas the estimates of traditional PLS and regression were expected to be attenuated. We moreover predicted that the attenuation would be strong because the reliability of the latent variable ξ_2 is only just acceptable.
 - The effect of ξ_1 on η_4 (γ_{14}) was zero and substantially affected by multicollinearity. We expected PLSc and covariance-based SEM to estimate this effect consistently, whereas the estimates of traditional PLS and regression were expected to be biased upward, making them prone to inflated Type I errors.

- The effect of ξ_1 on η_5 was small ($\gamma_{15} = 0.22$) and substantially affected by multicollinearity. We expect PLSc and covariance-based SEM to estimate this effect consistently, whereas the estimates of traditional PLS and regression were assumed to be biased downward, making them prone to inflated Type II errors.

We created both normal and non-normal data. To study the effects of non-normality, we used the customary Fleishman–Vale–Maurelli procedure (Fleishman 1978; Vale and Maurelli 1983), which means that the standard normal latent variables were replaced by well-chosen linear combinations of powers of standard normal variables, whose correlations are such that the new latent variables have the same correlations as the original latent variables. *Well-chosen* means that specified requirements concerning the non-normal skewness and (excess) kurtosis are satisfied. If the suggested transformations leave the independence between latent variables and idiosyncratic errors intact, the asymptotic robustness of normal theory statistics may apply and lead the researcher to mistakenly believe that normality is not an issue (see Hu et al. 1992). We followed the latter authors in that we simply rescaled the vector of indicators by multiplying each component by the same independent random factor. Hu et al. (1992) chose $\sqrt{3/\chi^2(5)}$, whose squared value has expectation 1. This approach deliberately destroys the independence between the latent and idiosyncratic variables but leaves Σ and the linear relationships (as well as the symmetry) undisturbed. The

kurtosis of the indicators increases by six. The same effects can be obtained by multiplying each component by a standard normal variable Z , which appears to yield representative samples for smaller sizes. Therefore, this approach was used. In addition, we performed multiplication by means of a positive scale factor $\sqrt{\text{abs}(Z) \cdot \sqrt{\pi/2}}$, whose squared value also has expectation 1. This multiplication increases the kurtosis by a lesser amount, namely, $3\pi/2 - 3 = 1.7124$.

In addition to the distribution of data, we varied the number of observations. Taking into account that PLS is typically applied if the sample size is rather small, we chose sample sizes of 100, 200, and 500 observations. Overall, the computational experiment consisted of six cells: two levels of data distribution \times three levels of sample size. We generated 1,000 datasets per cell.

We estimated the structural equation model for each dataset using full-information maximum likelihood, generalized least squares, weighted least squares, diagonally weighted least squares, unweighted least squares, regression on sum scores, traditional PLS (using the factor weighting scheme), and PLSc. All computations were conducted using R 3.0.2 (R Core Team 2013). We relied on our own implementation for traditional and consistent PLS, whereas we used the lavaan package, version 0.5-15 (Rosseel 2012), as an implementation of the five covariance-based SEM techniques. To avoid confounding differences in techniques and differences in parameterization (for a profound discussion on this matter, see Goodhue et al. 2012a; Marcoulides et al. 2012), we kept the parameterization across techniques as similar as possible (see Table 1). The only exception was the technique-induced difference to determine the way in which indicator error covariances would be treated: Whereas in covariance-based SEM the indicator error covariances are fixed to zero, the other techniques simply assume them to be zero, although they may ultimately turn out to be nonzero.

Results

The simulation study allowed us to compare the performance of PLSc, covariance-based SEM (FIML, GLS, WLS, DWLS, and ULS), traditional PLS, and a (multiple) regression on sum scores. A first issue that emerged relates to the convergence behavior of the respective techniques. Based on prior research (Henseler 2010; Jöreskog and Lawley 1968), we expected there to be substantially more cases of non-convergence for covariance-based SEM than for the other techniques.

As Table 2 shows, the simulation study only partly confirmed our expectation. Six out of the eight considered techniques,

among them PLSc, demonstrated an excellent convergence behavior and provided estimates more in than 99 percent of the simulation runs. In our computational experiment, FIML performed substantially better than in comparable studies (see Reinartz et al. 2009). Only two techniques stood out as negative: GLS and WLS. GLS faces a substantial amount of nonconvergence in the case of a small sample size, particularly if the data are not normally distributed. WLS is not executed at all for sample sizes smaller than 210 observations.⁵ For the remaining analyses, we only included those simulation runs in which a technique converged, which may mean a slightly reduced effective sample size in some experimental cells.

Second, we compared the eight techniques with regard to their consistency. We expected PLSc to provide consistent estimates, just like covariance-based SEM, but to clearly outperform traditional PLS and regression on sum scores. To empirically verify our expectation, we compared the techniques' raw biases. Raw bias (RB) is defined as follows:

$$\text{RB} = \frac{1}{T} \sum_{t=1}^T \hat{\theta}_t - \theta \quad (7)$$

in which θ is the true value of the model parameter of interest, $\hat{\theta}_t$ is its estimate in the t^{th} simulation run, and T is the number of simulation runs per experimental condition. We show the raw bias of each technique under every condition in Table 3.

Overall, we could confirm that the raw bias of PLSc is comparable to various covariance-based SEM techniques' raw biases. PLSc as well as the five covariance-based SEM techniques exhibited neither significant nor substantial biases for any of the six effects in the population model. Moreover, the raw bias of PLSc estimates and its standard deviation decreased, with an increase in sample size, which illustrates the consistency of PLSc. The same was true for the covariance-based SEM techniques. In contrast, for traditional PLS and regression on sum scores, such a decrease in raw bias could only be observed for the effect γ_{23} , which is neither affected by multicollinearity nor attenuation. The results were thus mainly driven by the consistency. The consistent techniques, PLSc and covariance-based SEM, formed a group of six techniques, whose estimates were distributed similarly. The expected values for all six effects were close to the true parameter values. The inconsistent techniques, traditional PLS and regression on sum scores, formed another group in which the expected values of traditional PLS estimates were

⁵WLS requires at least $N_{\min} = \frac{1}{2} v (v + 1)$ observations, where v is the number of observed variables. Our simulation model has 20 observed variables. Thus, N_{\min} is 210.

Table 1. Parameterization for the Simulation Study

Parameters	Consistent PLS	Traditional PLS	Regression on Sum Scores	Covariance-Based SEM
Path coefficients	free	free	free	free
Latent variable variances	fixed to 1	fixed to 1	fixed to 1	fixed to 1
Indicator weights	free	free	equality constraint	free
Indicator loadings	free	free	free	free
Indicator error variances	free	free	free	free
Disturbance term covariances	assumed to be 0	assumed to be 0	assumed to be 0	fixed to 0
Indicator error covariances	assumed to be 0	assumed to be 0	assumed to be 0	fixed to 0

Table 2. Occurrence of Non-Convergence

Distribution	Observations	PLSc	FIML	GLS	WLS	DWLS	ULS	PLS	OLS
Normal	100	0.0%	0.0%	2.3%	100.0%	0.2%	0.0%	0.0%	0.0%
	200	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%
	500	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Non-normal	100	0.1%	0.9%	12.6%	100.0%	0.9%	0.6%	0.1%	0.0%
	200	0.0%	0.0%	0.3%	100.0%	0.2%	0.0%	0.0%	0.0%
	500	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

Table 3. Raw Bias (Expected minus True Value) of PLSc, FIML, GLS, WLS, DWLS, ULS, PLS, and OLS

Distri- bution	Obser- vations	Tech- nique	$\xi_2 \rightarrow \eta_3$ $\gamma_{23} = 0.00$	$\eta_3 \rightarrow \eta_5$ $\beta_{35} = 0.35$	$\xi_2 \rightarrow \eta_4$ $\gamma_{24} = 0.70$	$\xi_2 \rightarrow \eta_5$ $\gamma_{25} = -0.70$	$\xi_1 \rightarrow \eta_4$ $\gamma_{14} = 0.00$	$\xi_1 \rightarrow \eta_5$ $\gamma_{15} = 0.22$
Normal	100	PLSc	-0.0111 (0.1644)	0.0054 (0.1158)	0.0231 (0.1971)	-0.0312 (0.2097)	-0.0184 (0.1992)	0.0139 (0.2033)
		FIML	0.0021 (0.1264)	-0.0024 (0.0957)	0.0212 (0.2090)	-0.0223 (0.2082)	-0.0277 (0.2116)	0.0210 (0.2058)
		GLS	0.0044 (0.1553)	0.0000 (0.1179)	0.0115 (0.2624)	-0.0185 (0.2643)	0.0056 (0.2617)	0.0012 (0.2596)
		WLS	—	—	—	—	—	—
		DWLS	-0.0055 (0.1324)	-0.0024 (0.0989)	0.0207 (0.2153)	-0.0359 (0.2205)	-0.0269 (0.2167)	0.0304 (0.2173)
		ULS	-0.0052 (0.1325)	-0.0013 (0.0986)	0.0237 (0.2098)	-0.0368 (0.2152)	-0.0281 (0.2116)	0.0303 (0.2116)
		PLS	-0.0087 (0.1283)	-0.0373 (0.0844)	-0.2736 (0.0930)	0.2432 (0.0980)	0.1574 (0.1052)	-0.2152 (0.1028)
		OLS	0.0019 (0.1027)	-0.0570 (0.0828)	-0.3049 (0.0962)	0.2700 (0.1013)	0.1657 (0.1073)	-0.2238 (0.1047)
	200	PLSc	-0.0089 (0.1073)	0.0029 (0.0730)	0.0071 (0.1349)	-0.0079 (0.1309)	-0.0033 (0.1321)	0.0017 (0.1309)
		FIML	-0.0014 (0.0854)	-0.0004 (0.0641)	0.0068 (0.1345)	-0.0052 (0.1257)	-0.0074 (0.1323)	0.0052 (0.1275)
		GLS	-0.0071 (0.0969)	-0.0004 (0.0704)	0.0041 (0.1443)	-0.0113 (0.1366)	0.0084 (0.1464)	-0.0077 (0.1383)
		WLS	—	—	—	—	—	—
		DWLS	-0.0048 (0.0868)	-0.0011 (0.0677)	0.0048 (0.1385)	-0.0089 (0.1376)	-0.0062 (0.1355)	0.0078 (0.1398)
		ULS	-0.0047 (0.0868)	-0.0007 (0.0676)	0.0071 (0.1371)	-0.0107 (0.1369)	-0.0075 (0.1345)	0.0091 (0.1395)
		PLS	-0.0068 (0.0837)	-0.0398 (0.0556)	-0.2814 (0.0661)	0.2500 (0.0687)	0.1623 (0.0693)	-0.2142 (0.0729)
		OLS	-0.0003 (0.0703)	-0.0548 (0.0569)	-0.3059 (0.0679)	0.2712 (0.0698)	0.1700 (0.0706)	-0.2239 (0.0746)
	500	PLSc	-0.0009 (0.0657)	0.0019 (0.0451)	0.0054 (0.0863)	-0.0087 (0.0799)	-0.0041 (0.0876)	0.0049 (0.0782)
		FIML	0.0013 (0.0549)	0.0000 (0.0395)	0.0047 (0.0853)	-0.0076 (0.0771)	-0.0053 (0.0862)	0.0064 (0.0763)
		GLS	-0.0024 (0.0574)	-0.0003 (0.0410)	0.0056 (0.0884)	-0.0109 (0.0786)	-0.0007 (0.0890)	0.0017 (0.0785)
		WLS	-0.0023 (0.0754)	0.0000 (0.0521)	0.0055 (0.1081)	-0.0127 (0.0970)	-0.0016 (0.1070)	0.0046 (0.0962)
		DWLS	0.0000 (0.0560)	-0.0002 (0.0423)	0.0041 (0.0874)	-0.0083 (0.0814)	-0.0046 (0.0886)	0.0064 (0.0806)
		ULS	0.0000 (0.0560)	0.0001 (0.0421)	0.0050 (0.0869)	-0.0089 (0.0813)	-0.0052 (0.0879)	0.0068 (0.0806)
		PLS	-0.0007 (0.0511)	-0.0421 (0.0350)	-0.2829 (0.0441)	0.2496 (0.0436)	0.1617 (0.0478)	-0.2117 (0.0461)
		OLS	0.0019 (0.0454)	-0.0535 (0.0360)	-0.3026 (0.0447)	0.2663 (0.0441)	0.1683 (0.0479)	-0.2216 (0.0462)

Table 3. Raw Bias (Expected minus True Value) of PLSc, FIML, GLS, WLS, DWLS, ULS, PLS, and OLS (Continued)

Distribution	Observations	Technique	$\xi_2 \rightarrow \eta_3$ $\gamma_{23} = 0.00$	$\eta_3 \rightarrow \eta_5$ $\beta_{35} = 0.35$	$\xi_2 \rightarrow \eta_4$ $\gamma_{24} = 0.70$	$\xi_2 \rightarrow \eta_5$ $\gamma_{25} = -0.70$	$\xi_1 \rightarrow \eta_4$ $\gamma_{14} = 0.00$	$\xi_1 \rightarrow \eta_5$ $\gamma_{15} = 0.22$
Non-normal	100	PLSc	-0.0209 (0.2154)	0.0195 (0.1692)	0.0234 (0.2586)	-0.0564 (0.3238)	-0.0195 (0.2455)	0.0352 (0.3007)
		FIML	-0.0012 (0.1513)	0.0046 (0.1173)	0.0310 (0.4889)	-0.0464 (0.3728)	-0.0450 (0.4869)	0.0500 (0.3661)
		GLS	0.0122 (0.2035)	0.0160 (0.1618)	-0.0013 (0.4672)	-0.0173 (0.5676)	0.0027 (0.4473)	0.0141 (0.5532)
		WLS	—	—	—	—	—	—
		DWLS	-0.0115 (0.1580)	0.0056 (0.1235)	0.0185 (0.2875)	-0.0551 (0.3367)	-0.0304 (0.2806)	0.0509 (0.3339)
		ULS	-0.0108 (0.1606)	0.0088 (0.1240)	0.0284 (0.2801)	-0.0577 (0.3246)	-0.0346 (0.2721)	0.0508 (0.3211)
		PLS	-0.0150 (0.1614)	-0.0243 (0.1041)	-0.2727 (0.1211)	0.2402 (0.1243)	0.1625 (0.1205)	-0.2143 (0.1207)
		OLS	-0.0008 (0.1270)	-0.0496 (0.1033)	-0.3088 (0.1238)	0.2718 (0.1241)	0.1689 (0.1217)	-0.2226 (0.1249)
	200	PLSc	-0.0029 (0.1381)	0.0091 (0.0944)	0.0096 (0.1724)	-0.0166 (0.1664)	-0.0036 (0.1690)	0.0082 (0.1584)
		FIML	0.0043 (0.1083)	-0.0013 (0.0819)	0.0054 (0.1828)	-0.0112 (0.1659)	-0.0093 (0.1779)	0.0152 (0.1621)
		GLS	-0.0003 (0.1260)	-0.0014 (0.0935)	0.0024 (0.1992)	-0.0160 (0.1770)	0.0142 (0.1966)	-0.0051 (0.1782)
		WLS	—	—	—	—	—	—
		DWLS	-0.0004 (0.1099)	-0.0009 (0.0860)	0.0034 (0.1838)	-0.0146 (0.1789)	-0.0065 (0.1784)	0.0161 (0.1741)
		ULS	0.0006 (0.1117)	0.0014 (0.0865)	0.0091 (0.1828)	-0.0206 (0.1782)	-0.0102 (0.1786)	0.0202 (0.1737)
		PLS	-0.0023 (0.1071)	-0.0370 (0.0701)	-0.2808 (0.0855)	0.2494 (0.0849)	0.1654 (0.0911)	-0.2145 (0.0874)
		OLS	0.0042 (0.0876)	-0.0551 (0.0727)	-0.3085 (0.0878)	0.2733 (0.0878)	0.1732 (0.0920)	-0.2226 (0.0905)
	500	PLSc	-0.0021 (0.0840)	0.0025 (0.0577)	0.0076 (0.1089)	-0.0074 (0.0954)	-0.0043 (0.1065)	0.0024 (0.0930)
		FIML	0.0032 (0.0687)	0.0002 (0.0505)	0.0070 (0.1090)	-0.0064 (0.0945)	-0.0070 (0.1056)	0.0053 (0.0940)
		GLS	-0.0010 (0.0729)	-0.0005 (0.0529)	0.0064 (0.1136)	-0.0083 (0.0962)	0.0022 (0.1107)	-0.0056 (0.0952)
		WLS	0.0003 (0.1004)	-0.0001 (0.0694)	0.0008 (0.1418)	-0.0154 (0.1312)	-0.0016 (0.1341)	0.0057 (0.1265)
		DWLS	0.0001 (0.0699)	-0.0004 (0.0541)	0.0054 (0.1125)	-0.0076 (0.1007)	-0.0059 (0.1097)	0.0063 (0.0997)
		ULS	0.0002 (0.0707)	0.0004 (0.0539)	0.0072 (0.1109)	-0.0097 (0.1000)	-0.0068 (0.1082)	0.0075 (0.0991)
		PLS	-0.0016 (0.0655)	-0.0413 (0.0451)	-0.2822 (0.0541)	0.2501 (0.0525)	0.1614 (0.0560)	-0.2130 (0.0554)
		OLS	0.0016 (0.0571)	-0.0541 (0.0463)	-0.3032 (0.0550)	0.2682 (0.0529)	0.1681 (0.0566)	-0.2224 (0.0557)

Note: Standard deviations in parentheses. Raw bias values in bold are significantly different from zero ($p < .05$, two-sided test).

slightly closer to the true parameters compared to regression on sum scores. The estimates of the effect of η_3 on η_5 were attenuated by traditional PLS and regression on sum scores, whereas PLSc and covariance-based SEM estimated this effect consistently. A similar picture emerged for the effects of ξ_2 on η_4 and of ξ_2 on η_5 . Again, PLSc and covariance-based SEM estimated these effects consistently. In these two instances, the attenuation encountered by traditional PLS and regression on sum scores was even more pronounced. PLSc and covariance-based SEM consistently estimated the effect of ξ_1 on η_4 as zero, whereas the estimates of traditional PLS and regression were biased upward, making them prone to inflated Type I errors. Finally, whereas PLSc and covariance-based SEM estimated the effect of ξ_1 on η_5 consistently, traditional PLS and regression tended to underestimate this effect, which rendered these two techniques prone to inflated Type II errors.

Because the occurrence of Type I and Type II errors is a pivotal problem of hypothesis testing, we examined the statistical inference provided by the eight techniques. Because

FIML is a parametric approach, we expected it to be more powerful, but also more dependent on normality than PLSc as a non-parametric approach. This outcome coincides with Jöreskog's (1983) finding that maximum likelihood estimates are more efficient than other covariance-based SEM techniques. Moreover, in line with Goodhue et al. (2012b), we did not expect strong differences between techniques if there was no multicollinearity. However, if multicollinearity was present, we expected that, in some instances, traditional PLS and regression on sum scores would exhibit an unacceptable number of false positives, whereas in other instances, they would show a substantially lower statistical power.

Table 4 provides an overview of the empirically achieved statistical power of each technique. For each of the experimental conditions (i.e., for each combination of sample size and data distribution), the numbers in the table represent the proportion of simulation runs for which a particular technique identified a certain effect as significant (using a significance level of $\alpha = 0.05$).

Table 4. Statistical Power of PLSc, FIML, GLS, WLS, DWLS, ULS, PLS, and OLS

Distribution	Observations	Technique	$\xi_2 \rightarrow \eta_3$ $\gamma_{23} = 0.00$	$\eta_3 \rightarrow \eta_5$ $\beta_{35} = 0.35$	$\xi_2 \rightarrow \eta_4$ $\gamma_{24} = 0.70$	$\xi_2 \rightarrow \eta_5$ $\gamma_{25} = -0.70$	$\xi_1 \rightarrow \eta_4$ $\gamma_{14} = 0.00$	$\xi_1 \rightarrow \eta_5$ $\gamma_{15} = 0.22$
Normal	100	PLSc	5.4%	56.8%	73.2%	71.4%	2.1%	16.1%
		FIML	2.0%	91.4%	73.7%	84.3%	4.6%	20.4%
		GLS	2.3%	76.9%	63.9%	74.9%	6.5%	15.0%
		WLS	—	—	—	—	—	—
		DWLS	51.8%	99.2%	83.7%	94.0%	1.2%	5.8%
		ULS	52.1%	99.2%	85.5%	95.2%	1.6%	8.2%
		PLS	5.4%	80.6%	89.3%	89.8%	33.3%	5.0%
		OLS	4.8%	91.0%	96.0%	98.4%	35.2%	5.3%
	200	PLSc	4.5%	95.8%	99.3%	99.5%	3.2%	42.9%
		FIML	2.9%	100.0%	99.5%	99.7%	3.7%	45.4%
		GLS	3.7%	99.6%	99.0%	99.7%	5.3%	37.1%
		WLS	—	—	—	—	—	—
		DWLS	51.2%	100.0%	99.6%	99.9%	1.3%	28.9%
		ULS	50.9%	100.0%	99.6%	99.9%	1.4%	36.3%
		PLS	4.5%	97.9%	99.9%	99.9%	61.2%	5.4%
		OLS	4.3%	99.8%	99.9%	100.0%	64.2%	5.9%
	500	PLSc	5.1%	100.0%	100.0%	100.0%	4.4%	83.9%
		FIML	4.1%	100.0%	100.0%	100.0%	5.1%	87.2%
		GLS	4.4%	100.0%	100.0%	100.0%	6.1%	83.3%
		WLS	24.3%	100.0%	100.0%	100.0%	21.4%	87.5%
		DWLS	54.7%	100.0%	100.0%	100.0%	2.0%	85.0%
		ULS	54.1%	100.0%	100.0%	100.0%	2.4%	87.7%
		PLS	5.1%	100.0%	100.0%	100.0%	92.5%	5.5%
		OLS	4.9%	100.0%	100.0%	100.0%	94.9%	5.2%
Non-normal	100	PLSc	5.0%	24.9%	34.4%	24.9%	1.2%	3.7%
		FIML	4.9%	87.2%	63.9%	78.6%	8.9%	25.0%
		GLS	7.8%	70.9%	51.3%	61.0%	10.4%	19.4%
		WLS	—	—	—	—	—	—
		DWLS	48.3%	94.3%	56.3%	74.6%	0.6%	1.7%
		ULS	58.2%	96.6%	76.7%	89.2%	2.9%	12.3%
		PLS	5.0%	62.6%	69.5%	70.9%	25.4%	4.4%
		OLS	12.4%	87.7%	91.9%	94.9%	39.2%	10.3%
	200	PLSc	5.2%	74.8%	82.2%	80.5%	2.3%	19.3%
		FIML	8.2%	98.5%	96.7%	99.4%	11.0%	46.4%
		GLS	9.9%	96.8%	94.2%	97.9%	12.5%	35.3%
		WLS	—	—	—	—	—	—
		DWLS	49.8%	99.7%	93.0%	98.2%	1.8%	7.9%
		ULS	60.2%	99.9%	98.5%	99.7%	6.1%	37.7%
		PLS	5.2%	90.6%	95.5%	95.6%	46.8%	5.6%
		OLS	11.0%	98.7%	99.7%	99.9%	60.5%	11.2%
	500	PLSc	6.7%	99.8%	99.8%	99.8%	3.5%	61.5%
		FIML	10.9%	100.0%	100.0%	100.0%	11.3%	81.3%
		GLS	11.7%	100.0%	100.0%	100.0%	11.7%	76.8%
		WLS	25.1%	99.9%	99.8%	100.0%	19.8%	71.9%
		DWLS	53.1%	100.0%	99.9%	100.0%	1.2%	57.1%
		ULS	61.7%	100.0%	100.0%	100.0%	6.3%	83.8%
		PLS	6.7%	99.8%	100.0%	100.0%	80.7%	5.5%
		OLS	13.0%	100.0%	100.0%	100.0%	91.0%	9.7%

Our analysis began with the situation of normal-distributed data. If no multicollinearity was present, five of the eight techniques maintained an acceptable level of Type I error for a population effect of zero (γ_{23}). In contrast, ULS and DWLS found a significant effect in more than half of all simulation runs, and WLS still found a significant effect in almost a quarter of the simulation runs.⁶ If there was a medium effect in the population model (as with β_{35}), all techniques were able to detect the effect in the majority of cases. For medium to large sample sizes, the statistical power of all the techniques was close to 100 percent. However, for small sample sizes, the statistical power decreased, and, in particular, PLSc had a somewhat lower statistical power than the other techniques. A similar result was found for strong effects under high levels of multicollinearity (as with γ_{24} and γ_{25}). Here, PLSc was comparable to GLS, and almost as powerful as FIML.

The most important insights referred to small or zero effects under multicollinearity. We found that neither PLS nor regression was capable of detecting a small effect (γ_{15}), regardless of the sample size. In contrast, although PLSc and the covariance-based SEM techniques had a low statistical power for small sample sizes, the statistical power reached 80 percent in the case of 500 observations. In other words, the inconsistent techniques of PLS and regression evoke an undue number of Type II errors, whereas the consistent techniques do not. A different phenomenon was observed for the zero effect under the multicollinearity condition (γ_{14}). Whereas PLSc and the covariance-based SEM techniques (except WLS) maintained acceptable levels of Type I error, PLS and regression found significant effects in more than 30 percent of the cases for small sample sizes and, in more than 90 percent of the cases, for large sample sizes. In other words, the inconsistent techniques, traditional PLS and regression on sum scores, evoked an undue number of Type I errors. Moreover, WLS showed an increased level of Type I errors.

For non-normal data, the general pattern remained similar, but there were also some important differences compared to normal data. The parametric techniques—covariance-based SEM and regression—faced an increase in statistical power. Although this did not appear to be problematic for effects that exist in the population model, it had detrimental outcomes for effects that were essentially zero. For instance, for the effect γ_{23} in the case of 500 observations, the Type I error of all covariance-based SEM techniques exceeded 10 percent. In contrast, PLSc did not face this problem. Instead, PLSc maintained acceptable levels of Type I error under all conditions.

⁶We cannot completely rule out the possibility that this finding is due to problems in the software rather than the techniques themselves. Although different software provides the same parameter estimates, we could not find other software implementations providing inference for standardized effects using ULS, DWLS, or WLS with continuous data.

Nevertheless, this came at a cost: Among the consistent techniques, PLSc typically had the lowest statistical power when it came to detecting effects that exist in the population model. However, the size of the difference in power between techniques differed among paths. While the difference was substantial in some instances, it was negligible in others.

An Example from the Information Systems Domain

We present an example from the IS domain to demonstrate that choosing one technique over another can alter the conclusions drawn from an empirical analysis. We reanalyzed the data that Chin et al. (2003) initially used to demonstrate an extension to PLS.⁷ Chin et al.'s paper, which was published in the journal *Information Systems Research*, has found ample dissemination, as is evidenced by its more than 2,700 citations on Google Scholar as of December 2014. Figure 5 depicts the model that we reestimated. It is a small technology acceptance model in which the perceived exogenous constructs of usefulness and enjoyment are used to explain the intention to regularly use electronic mail.

We estimated the two path coefficients of this model using traditional PLS as well as PLSc. Moreover, we applied bootstrapping with 10,000 bootstrap samples and determined the 95 percent bootstrap percentile confidence intervals. Table 5 shows the results that we obtained using the two techniques. With regard to the two path coefficients, the differences between the estimates were rather small. The explanation for the low degree of differences is obvious: Because the reliabilities estimated by ρ_A were relatively high (0.8236 for enjoyment) or even very high (0.9720 for perceived usefulness and 0.9473 for intention to use), the correction undertaken by PLSc was weak. Unsurprisingly, the two techniques have led researchers to the same conclusion: perceived usefulness and enjoyment have a positive and significant impact on the intention to regularly use electronic mail. In the case of the first effect, one can also see that the correction for attenuation through PLSc does not necessarily imply increased estimates but can also lead to decreased estimates.

A small difference between the results of PLS and PLSc became apparent when we looked at the coefficients of determination: the R^2 was 0.5162 for PLSc versus 0.4650 for traditional PLS. Whereas PLSc could thus explain more than half of the variation in the dependent variable, the majority of variance remained unexplained when traditional PLS was used.

⁷We are grateful to Wynne W. Chin for providing us with the data.

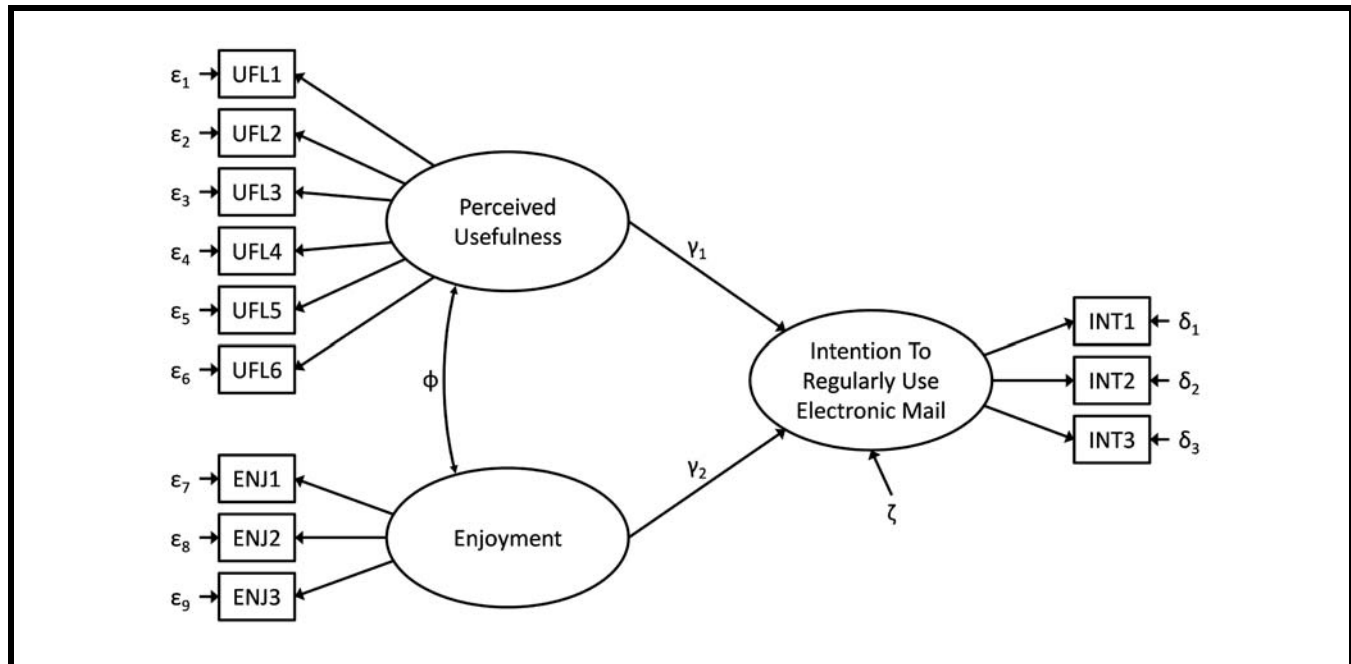


Figure 5. An Example Model from Information Systems Research (adapted from Chin et al. 2003; main effects only)

Table 5. Results of PLS and PLSc for the Example Model

Hypothesis	Technique	Estimate	95% CI	Conclusion
$\gamma_1 \neq 0$	PLS	0.5165	[0.3912; 0.6329]	supported
	PLSc	0.5069	[0.3212; 0.7477]	supported
$\gamma_2 \neq 0$	PLS	0.2689	[0.1725; 0.3688]	supported
	PLSc	0.3130	[0.1708; 0.4481]	supported
$\gamma_1 - \gamma_2 \neq 0$	PLS	0.2476	[0.0442; 0.4454]	supported
	PLSc	0.1940	[-0.0750; 0.5480]	not supported

In addition to this subtle difference in R^2 , we observed that, depending on the method, a researcher might come to substantially different conclusions. A legitimate question that managers, scientists, policymakers, or other interested parties may ask is, “Does perceived usefulness have a stronger effect on usage intention than enjoyment?” This question can be answered by means of a hypothesis test for the difference between the two effects. The output of such a hypothesis test is presented at the bottom of Table 5. In this case, it did matter which technique was used. Whereas PLSc could not confirm a significantly stronger influence of perceived usefulness, traditional PLS could.

In contrast to the simulation study presented earlier, we do not know the true underlying model for this example. We only

know that if the common factor model is correct for all three constructs (which needs to be shown), we can expect the estimates of PLSc, but not of PLS, to be consistent.

Implications and Guidelines

Selecting the Optimal Method of Analysis

As our study has illustrated, traditional PLS or regression on sum scores provide inconsistent estimates if the common factor model holds for all constructs. This inconsistency of estimates can lead to unfavorable consequences for hypothesis testing, namely increased Type I and Type II errors.

Therefore, for the purpose of hypothesis testing, we recommend that, whenever the common factor model holds, researchers should prefer PLSc or covariance-based SEM over regression on sum scores or traditional PLS. Conversely, if the common factor model is wrong but the composite model holds, traditional PLS should be the method of choice (Dijkstra 2013; Henseler et al. 2014). We anticipate that structural equation models that contain both composite and common factor models will become the dominant domain of PLS path modeling (see also Sarstedt, Ringle, and Hair 2014). Traditional PLS will provide all construct scores for these types of structural equation models, and the correction for attenuation using the new reliability coefficient ρ_A will be applied to the reflective constructs.

As a rule of thumb, one could say that to avoid inflated Type I and Type II errors, one should always choose a statistical method that can consistently estimate the measurement models at hand. Specifically, if the composite model is true, one should use a method that can address composites, such as PLS or regression based on sum scores. If the common factor model is true, one should use a method that can consistently estimate these types of models, such as CBSEM or PLSc. If a structural equation model contains both composites and common factors, PLSc (correcting only those constructs that are reflective) is the method of choice.

Occasionally, covariance-based SEM does not converge and thus cannot provide any estimates (Reinartz et al. 2009). Researchers usually solve this problem by changing the model (respecification) or the data (e.g., by transforming variables or eliminating influential observations). Our paper offers another solution: the use of PLSc, which can be applied without having to change the model or the data. Owing to its favorable convergence behavior, PLSc should be the method of choice if the same model is to be estimated many times using different datasets and non-convergence is not a defensible option. Typical examples are technology acceptance models (e.g., Davis 1989) or national customer satisfaction indices (e.g., Fornell 1992).

The distribution of data is another aspect to consider. Although the degree of non-normality induced in our simulation was rather modest, that is, the data were not skewed and the kurtosis was in an acceptable range (see Cameron 2003; Kline 2010), the performance of the various covariance-based SEM techniques dropped substantially. First, the convergence behavior of GLS and, to a smaller extent, FIML, DWLS, and ULS, deteriorated. Second, the operating characteristics of covariance-based SEM changed substantially, showing an unacceptably high rate of Type I errors. In contrast, a non-normal distribution of data hardly affected PLSc. The Type I error of PLSc was not a point of concern, and cases of non-

convergence were unlikely. Based on our findings, we therefore recommend that researchers use PLSc if they are unsure of whether their data are normally distributed.

If researchers analyze a sample that is generalized to the population, a method's statistical power becomes relevant. PLSc's statistical power is comparable to that of GLS, but slightly lower than that of FIML, meaning that PLSc requires more observations than FIML to render an effect significant.

There are also situations in which analysts should hesitate to use PLSc. The iterative PLS algorithm determines the indicator weights of a proxy relying on the nomological net of the respective construct. If there is no nomological net, which means that a construct is perfectly unrelated to all other constructs in the model, PLS and, consequently, PLSc will yield large standard errors and will thus be inefficient. If researchers cannot rely on prior research to provide a sufficiently strong nomological net, they should consider applying Cohen's (1982) set correlation. This tool can be of high diagnostic value because it can establish confidence in that blocks of indicators can be expected to correlate. If a model is small (for instance, if it involves only two latent variables), researchers should only employ PLSc if they are confident that the latent variables are interrelated. Moreover, PLSc should not be used to demonstrate that a latent variable is unrelated to all other latent variables in a model.

Even if the common factor model holds, there may be instances in which analysts would prefer traditional PLS over covariance-based SEM or PLSc, such as in purely prediction-oriented research (Sarstedt, Ringle et al. 2014). As Becker, Rai, and Rigdon (2013) put it,

Composite-based methods like partial least squares (PLS) path modeling have an advantage over factor-based methods (like CB-SEM) because they yield determinate predictions, while factor-based methods' prediction is constrained in this regard by factor indeterminacy. To maximize practical relevance, research findings should extend beyond the study's own data.

Finally, if certain misspecifications are present in the model, limited-information approaches, such as PLSc, have an advantage over full-information approaches. This occurs

because if there is a misspecification in one part of the model and if the model is quite complicated with many equations, this misspecification will not bias estimates in other parts of the model as would full-information estimators (Antonakis et al. 2010, p. 1102).

For instance, structural equations estimated using PLS_c remain unaffected by the measurement model misspecification of constructs that do not form part of the respective equation or by the misspecification of other structural equations.

Broadening the Spectrum of Structural Equation Modeling Techniques

In this paper, we compared essentially eight structural equation modeling techniques. These techniques included PLS_c and five full-information covariance-based SEM techniques as estimators of common factor models on the one hand, and traditional PLS and regression on sum scores as estimators of composite models on the other hand. An array of other structural equation modeling techniques are available, and researchers should consider using them if they prove more appropriate for meeting their research objective. We thereby concur with Hair, Ringle, and Sarstedt (2012), who “call for a more balanced and informed judgment of the available SEM methods” (p. 314).

In addition to the full-information approaches to covariance-based SEM, there are also numerous limited-information approaches to structural equation modeling that use common factor models for the measurement part. For instance, in their regression component analysis, Schönemann and Steiger (1976) first estimated the factor loadings and then determined error variances from what was left over. However, in contrast to PLS_c, regression component analysis requires the application of factor model software, whereas PLS_c determines the loadings based on the PLS weight estimates. Moreover, Hägglund (1982) introduced an instrumental variable approach to factor analysis, and Jöreskog (1983) described a two-stage least squares method. Hägglund's and Jöreskog's approaches yield consistent estimates. Two-stage least squares, as an estimator of structural equation models, has weaker distributional assumptions and is computationally less demanding than full-information covariance-based SEM techniques (Bollen 1996).

There are many variance-based structural equation modeling techniques to choose from other than regression on sum scores and PLS. As an alternative to sum scores, principal component and common factor analysis offer several options for creating scores. If researchers want to calibrate the indicator weights based on the nomological net of the latent variables, several alternatives are available. For instance, besides PLS, there are best fitting proper indices (Dijkstra and Henseler 2011), generalized structured component analysis (Henseler 2012; Hwang et al. 2004), regularized generalized canonical correlation analysis (Tenenhaus and Tenenhaus 2011), and various forms of creating generalized canonical variables (Kettenring 1971).

Ex Post Modification of Traditional PLS Results

It has become evident that the estimates of traditional PLS as well as of other variance-based structural equation modeling techniques, such as regression on sum scores or generalized structured component analysis, are not consistent if the common factor model holds for the construct measurement. The lack of consistency means that meta-analyses based on these techniques are not guaranteed to come closer to the true value than are single studies. It also means that researchers may have come to wrong conclusions in more cases than anticipated, despite having carefully selected adequate alpha levels and statistical power.⁸

At least for PLS, the advent of the new reliability coefficient ρ_A has created the possibility of calculating consistent estimates *ex post*. To do so, analysts require the inter-construct correlations, the inter-indicator correlations, and the indicator weights. Because inter-indicator correlations and indicator weights for reflective constructs do not belong to the common catalogs of reporting standards (Chin 2010; Gefen et al. 2011; Hair, Sarstedt et al. 2012; Henseler et al. 2009; Ringle et al. 2012), analysts should follow a two-step approach. In a first step, they should verify whether a correction for attenuation using the reported reliability measure(s) leads to substantially different empirical findings. If it does, they should, in a second step, request the inter-indicator correlations and the indicator weights from the authors to determine ρ_A . Once ρ_A is available, analysts can apply it to obtain consistent inter-construct correlations, which they can use to estimate consistent path coefficients. That is, the last two steps of PLS_c can be performed.

Conclusions and Future Research

PLS can be viewed from two basic perspectives. On the one hand, it can be regarded as an extension of principal components and canonical variables, as Herman Wold has maintained on various occasions (see any of his works). This view has also been maintained in more recent work (Dijkstra and Henseler 2011; Rigdon 2012). The alternative view, which Wold has also held at times, is that PLS is a distribution-free way to consistently estimate covariance structures or second-order factor models. In light of the findings presented in this paper, this view is less tenable or productive. Of course, the

⁸An anonymous reviewer noted that the large number of false positives in PLS and OLS obtained in our simulation study (see Table 4) should not be used as a simple gauge for the proportion of studies with likely erroneous conclusions. The absolute correlations typically observed in IS research are usually much smaller than those in our simulation study and so will be the number of false positives. We thank the reviewer for this note.

discrepancies between the probability limits of the PLS estimators, as induced by any of the modes, and the true parameter values lean toward zero as the quality of the proxies lean toward one (consistency at large). However, in actual applications, the quality is never maximal, and elements such as path coefficients or construct loadings tend to be sensitive to deviations from the ideal. Therefore, if one wants to use PLS as a consistent and numerically efficient alternative to any of the well-established methods for estimating covariance structures, it seems imperative to use PLSc.

PLSc builds on and enhances traditional PLS in several ways. First, as with traditional PLS, PLSc uses the iterative PLS algorithm to obtain construct scores. Second, unlike traditional PLS, PLSc employs a correction factor to obtain consistent construct correlations and consistent indicator loadings if the common factor model holds true. Third, both traditional PLS and PLSc typically apply ordinary least squares to estimate the relationships between constructs according to the model specification. In principle, PLSc is readily equipped to go beyond the limiting realm of ordinary least squares and to use two-stage least squares, seemingly unrelated regression or other appropriate estimators for the structural model (an approach not entirely new to PLS path modeling; see Boardman et al. 1981). This would allow for analyzing non-recursive models and considering endogeneity in the structural model. Future research should try to clarify the performance of PLSc in combination with alternative estimators for the structural model.

PLS has traditionally been understood as a technique that is applicable when dealing with small sample sizes. The logic behind this understanding is that, regardless of whether the data at hand represent a sample or a population, the number of observations must be sufficiently high so that the method can technically produce estimates. In contrast to covariance-based SEM, PLS can even generate parameter estimates if the number of model parameters or the number of variables exceeds the number of observations (Henseler et al. 2014). None of the four steps of PLSc postulates a larger sample size for consistent PLS than for traditional PLS. Nevertheless, the small sample behavior of PLSc was not within the scope of the present paper and warrants a thorough investigation, particularly in light of the ongoing discussion on PLS' small sample properties (see Goodhue et al. 2013).

Similar to traditional PLS, PLSc relies on the nomological net to calibrate the measurement model (Rai et al. 2013). If a construct is orthogonal to all other constructs in a structural equation model, PLS (and thus PLSc) is not likely to outperform less sophisticated techniques, such as regressions, based on sum scores or principal components. One way to ensure the proper functioning of PLS is to include other constructs in the model (Henseler et al. 2014). Future research

should systematically explore the extent to which the efficiency of PLSc depends on the strength of the nomological net, and it should try to formulate minimum requirements for the nomological net. Moreover, such research should spell out recommendations regarding when to use PLSc or alternative SEM techniques.

A remaining weakness of PLSc is its somewhat lower power compared to covariance-based SEM. If deemed necessary, Bentler and Dijkstra's (1986) suggested approach can be used to overcome the somewhat lower statistical power of PLSc. Updating a consistent estimator once using the Newton-Raphson or Gauss-Newton methods for the first-order conditions of a suitably fitting function will result in an asymptotically efficient estimator. Therefore, PLS can be made as efficient as ML-estimators by introducing a simple adjustment to it followed by an iteration of a standard optimization routine (Bentler and Hwang 2014). However, it should be noted that the efficient approach leans heavily on the assumption that the model is correct. One may wish to stick to PLSc due to its potential robustness in the case of a (suspected) misspecification.

Traditional PLS already has a broad scope and range of applications as a multiple indicator approach to modeling conceptual variables that uses composites rather than common factors as proxies. PLSc broadens this scope by making the PLS family of methods the primary choice for structural equation models containing both common factors and composites. In principle, PLSc is not limited to common factors or composites as measurement models. There are more sophisticated measurement models that go beyond the simple factor/component dichotomy (Rigdon 2014). The management literature contains examples such as differences (e.g., Gundlach et al. 1995) and interactions (e.g., Gibson and Birkinshaw 2004) between indicators. Modeling more sophisticated measurement models by means of PLSc is well within reach.

The advent of PLSc has substantial consequences for methodological research in MIS. Not only does MIS research make ample use of PLS as a method of analysis, but also many extensions and advances of PLS can be credited to MIS researchers. Major examples are PLS-related reporting standards (Gefen et al. 2011) and PLS-based approaches to multigroup analysis (Chin and Dibbern 2010; Keil et al. 2000; Qureshi and Compeau 2009), testing moderating effects (Chin et al. 2003; Goodhue et al. 2007; Henseler and Chin 2010), detecting unobserved heterogeneity (Becker, Rai et al. 2013), assessing common method bias (Chin et al. 2012; Liang et al. 2007), testing for measurement invariance (Hsieh et al. 2008), modeling nonlinear relationships (Henseler et al. 2012), and analyzing hierarchical component models (Ringle et al. 2012; Wetzels et al. 2009). These examples have all been devel-

oped, discussed, or improved within the MIS domain. All of these extensions and advances must be examined with respect to their applicability to PLS. Several of them most likely need to be adapted to fully exploit the benefits of consistent PLS.

References

- Antonakis, J., Bendahan, S., Jacquart, P., and Lalive, R. 2010. "On Making Causal Claims: A Review and Recommendations," *The Leadership Quarterly* (21:6), pp. 1086-1120.
- Becker, J.-M., Rai, A., and Rigdon, E. E. 2013. "Predictive Validity and Formative Measurement in Structural Equation Modeling: Embracing Practical Relevance," in *Proceedings of the 34th International Conference on Information Systems*, Milan, Italy, December 15-18.
- Becker, J.-M., Rai, A., Ringle, C. M., and Völckner, F. 2013. "Discovering Unobserved Heterogeneity in Structural Equation Models to Avert Validity Threats," *MIS Quarterly* (37:3), pp. 665-694.
- Bentler, P. M., and Dijkstra, T. K. 1986. "Efficient Estimation Via Linearization in Structural Models," in *Multivariate analysis VI*, P. R. Krishnaiah (ed.), Amsterdam: North-Holland, pp. 9-43.
- Bentler, P. M., and Huang, W. 2014. "On Components, Latent Variables, PLS and Simple Methods: Reactions to Rigdon's Rethinking of PLS," *Long Range Planning* (47:3), pp. 138-145.
- Boardman, A. E., Hui, B. S., and Wold, H. 1981. "The Partial Least Squares-Fix Point Method of Estimating Interdependent Systems with Latent Variables," *Communications in Statistics: Theory and Methods* (10:7), pp. 613-639.
- Bollen, K. A. 1996. "An Alternative Two Stage Least Squares (2SLS) Estimator for Latent Variable Equations," *Psychometrika* (61:1), pp. 109-121.
- Cameron, A. C. 2003. "Kurtosis," in *The Sage Encyclopedia of Social Science Research Methods*, M. S. Lewis-Beck, A. Bryman, and T. F. Liao (eds.), Thousand Oaks, CA: Sage, pp. 543-544.
- Chin, W. W. 1998a. "Issues and Opinion on Structural Equation Modeling," *MIS Quarterly* (22:1), pp. 7-16.
- Chin, W. W. 1998b. "The Partial Least Squares Approach for Structural Equation Modeling," in *Modern Methods for Business Research*, G. A. Marcoulides (ed.), Mahwah, NJ: Lawrence Erlbaum Associates, pp. 295-336.
- Chin, W. W. 2010. "How to Write Up and Report PLS Analyses," in *Handbook of Partial Least Squares: Concepts, Methods and Applications*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang (eds.), New York: Springer, pp. 655-690.
- Chin, W. W., and Dibbern, J. 2010. "An Introduction to a Permutation-Based Procedure for Multi-Group PLS Analysis: Results of Tests of Differences on Simulated Data and a Cross-Cultural Analysis of the Sourcing of Information System Services between Germany and the USA," in *Handbook of Partial Least Squares: Concepts, Methods and Applications*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang (eds.), New York: Springer, pp. 171-193.
- Chin, W. W., Marcolin, B. L., and Newsted, P. R. 2003. "A Partial Least Squares Latent Variable Modeling Approach for Measuring Interaction Effects: Results from a Monte Carlo Simulation Study and an Electronic-Mail Emotion/Adoption Study," *Information Systems Research* (14:2), pp. 189-217.
- Chin, W. W., and Newsted, P. R. 1999. "Structural Equation Modeling Analysis with Small Samples Using Partial Least Squares," in *Statistical Strategies for Small Sample Research*, R. H. Hoyle (ed.), Thousand Oaks, CA: Sage Publications, pp. 307-341.
- Chin, W. W., Thatcher, J. B., and Wright, R. T. 2012. "Assessing Common Method Bias: Problems with the ULMC Technique," *MIS Quarterly* (36:3), pp. 1003-1019.
- Cohen, J. 1982. "Set Correlation as a General Multivariate Data-Analytic Method," *Multivariate Behavioral Research* (17:3), pp. 301-341.
- Cohen, J. 1988. *Statistical Power Analysis for the Behavioral Sciences*, Mahwah, NJ: Lawrence Erlbaum Associates.
- Cronbach, L. J. 1951. "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* (16:3), pp. 297-334.
- Davis, F. D. 1989. "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly* (13:3), pp. 319-340.
- Dijkstra, T. K. 1981. *Latent Variables in Linear Stochastic Models: Reflections on "Maximum Likelihood" and "Partial Least Squares" Methods*, Ph.D. Thesis, Groningen University, Groningen, The Netherlands.
- Dijkstra, T. K. 1983. "Some Comments on Maximum Likelihood and Partial Least Squares Methods," *Journal of Econometrics* (22:1-2), pp. 67-90.
- Dijkstra, T. K. 2010. "Latent Variables and Indices: Herman Wold's Basic Design and Partial Least Squares," in *Handbook of Partial Least Squares: Concepts, Methods and Applications*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang (eds.), New York: Springer, pp. 23-46.
- Dijkstra, T. K. 2011. "Consistent Partial Least Squares Estimators for Linear and Polynomial Factor Models," working paper, University of Groningen, Groningen, The Netherlands (<http://www.rug.nl/staff/t.k.dijkstra/research>).
- Dijkstra, T. K. 2013. "Composites as Factors," working paper, University of Groningen, Groningen, The Netherlands (<http://www.rug.nl/staff/t.k.dijkstra/composites-as-factors.pdf>).
- Dijkstra, T. K. 2014. "PLS' Janus Face—Response to Professor Rigdon's 'Rethinking Partial Least Squares Modeling: In Praise of Simple Methods,'" *Long Range Planning* (47:3), pp. 146-153.
- Dijkstra, T. K., and Henseler, J. 2011. "Linear Indices in Nonlinear Structural Equation Models: Best Fitting Proper Indices and Other Composites," *Quality & Quantity* (45:6), pp. 1505-1518.
- Dijkstra, T. K., and Henseler, J. 2015. "Consistent and Asymptotically Normal PLS Estimators for Linear Structural Equations," *Computational Statistics & Data Analysis* (81:1), pp. 10-23.
- Fleishman, A. I. 1978. "A Method for Stimulating Non-Normal Distributions," *Psychometrika* (43:4), pp. 521-532.
- Fornell, C. 1992. "A National Customer Satisfaction Barometer: The Swedish Experience," *Journal of Marketing* (56:1), pp. 6-21.
- Fornell, C., and Bookstein, F. L. 1982. "Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory," *Journal of Marketing Research* (19:4), pp. 440-452.

- Gefen, D., Straub, D. W., and Rigdon, E. E. 2011. "An Update and Extension to SEM Guidelines for Administrative and Social Science Research," *MIS Quarterly* (35:2), pp. iii-xiv.
- Gerbing, D. W., and Anderson, J. C. 1988. "An Updated Paradigm for Scale Development Incorporating Unidimensionality and its Assessment," *Journal of Marketing Research* (25:2), pp. 186-192.
- Gibson, C. B., and Birkinshaw, J. 2004. "The Antecedents, Consequences, and Mediating Role of Organizational Ambidexterity," *Academy of Management Journal* (47:2), pp. 209-226.
- Goodhue, D. L., Lewis, W., and Thompson, R. 2006. "PLS, Small Sample Size, and Statistical Power in MIS Research," in *Proceedings of the 39th Annual Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society Press, pp. 202b-202b.
- Goodhue, D. L., Lewis, W., and Thompson, R. 2007. "Statistical Power in Analyzing Interaction Effects: Questioning the Advantage of PLS with Product Indicators," *Information Systems Research* (18:2), pp. 211-227.
- Goodhue, D. L., Lewis, W., and Thompson, R. L. 2011. "A Dangerous Blind Spot in IS Research: False Positives Due to Multicollinearity Combined With Measurement Error," in *Proceedings of the 17th Americas Conference on Information Systems*, Detroit, MI.
- Goodhue, D. L., Lewis, W., and Thompson, R. 2012a. "Comparing PLS to Regression and LISREL: A Response to Marcoulides, Chin, and Saunders," *MIS Quarterly* (36:3), pp. 703-716.
- Goodhue, D. L., Lewis, W., and Thompson, R. 2012b. "Does PLS Have Advantages for Small Sample Size or Non-Normal Data?," *MIS Quarterly* (36:3), pp. 981-1001.
- Goodhue, D. L., Lewis, W., and Thompson, R. 2013. "Why You Shouldn't Use PLS: Four Reasons to Be Uneasy About Using PLS In Analyzing Path Models," in *Proceedings of the 46th Annual Hawaii International Conference on System Sciences*, Los Alamitos, CA: IEEE Computer Society Press, pp. 4739-4748.
- Gundlach, G. T., Achrol, R. S., and Mentzer, J. T. 1995. "The Structure of Commitment in Exchange," *Journal of Marketing* (59:1), pp. 78-92.
- Hägglund, G. 1982. "Factor Analysis by Instrumental Variables Methods," *Psychometrika* (47:2), pp. 209-222.
- Hair, J. F., Ringle, C. M., and Sarstedt, M. 2011. "PLS-SEM: Indeed a Silver Bullet," *Journal of Marketing Theory and Practice* (18:2), pp. 139-152.
- Hair, J. F., Ringle, C. M., and Sarstedt, M. 2012. "Partial Least Squares: The Better Approach to Structural Equation Modeling?," *Long Range Planning* (45:5-6), pp. 312-319.
- Hair, J. F., Sarstedt, M., Ringle, C. M., and Mena, J. A. 2012. "An Assessment of the Use of Partial Least Squares Structural Equation Modeling in Marketing Research," *Journal of the Academy of Marketing Science* (40:3), pp. 414-433.
- Heise, D. R., and Bohrnstedt, G. W. 1970. "Validity, Invalidity, and Reliability," in *Sociological Methodology*, E. F. Borgatta and G. W. Bohrnstedt (eds.), San Francisco: Jossey-Bass, pp. 104-129.
- Henseler, J. 2010. "On the Convergence of the Partial Least Squares Path Modeling Algorithm," *Computational Statistics* (25:1), pp. 107-120.
- Henseler, J. 2012. "Why Generalized Structured Component Analysis Is Not Universally Preferable to Structural Equation Modeling," *Journal of the Academy of Marketing Science* (40:3), pp. 402-413.
- Henseler, J., and Chin, W. W. 2010. "A Comparison of Approaches for the Analysis of Interaction Effects Between Latent Variables Using Partial Least Squares Path Modeling," *Structural Equation Modeling* (17:1), pp. 82-109.
- Henseler, J., Dijkstra, T. K., Sarstedt, M., Ringle, C. M., Diamantopoulos, A., Straub, D. W., Ketchen, Jr., D. J., Hair, J. F., Hult, G. T. M., and Calantone, R. J. 2014. "Common Beliefs and Reality about PLS: Comments on Rönkkö and Evermann (2013)," *Organizational Research Methods* (17:2), pp. 182-209.
- Henseler, J., Fassott, G., Dijkstra, T. K., and Wilson, B. 2012. "Analysing Quadratic Effects of Formative Constructs by Means of Variance-Based Structural Equation Modelling," *European Journal of Information Systems* (21:1), pp. 99-112.
- Henseler, J., Ringle, C. M., and Sarstedt, M. 2014. "A New Criterion for Assessing Discriminant Validity in Variance-Based Structural Equation Modeling," *Journal of the Academy of Marketing Science* (doi:10.1007/s11747-014-0403-8).
- Henseler, J., Ringle, C. M., and Sinkovics, R. R. 2009. "The Use of Partial Least Squares Path Modeling in International Marketing," in *New Challenges to International Marketing*, R. R. Sinkovics and P. N. Ghauri (eds.), Bingley, UK: Emerald Group Publishing, pp. 277-320.
- Hsieh, J. J. P.-A., Rai, A., and Keil, M. 2008. "Understanding Digital Inequality: Comparing Continued Use Behavioral Models of the Socio-Economically Advantaged and Disadvantaged," *MIS Quarterly* (32:1), pp. 97-126.
- Hu, L.-T., Bentler, P. M., and Kano, Y. 1992. "Can Test Statistics in Covariance Structure Analysis be Trusted?," *Psychological Bulletin* (112:2), pp. 351-362.
- Hwang, H., and Takane, Y. 2004. "Generalized Structured Component Analysis," *Psychometrika* (69:1), pp. 81-99.
- Jarvis, C. B., MacKenzie, S. B., and Podsakoff, P. M. 2003. "A Critical Review of Construct Indicators and Measurement Model Misspecification in Marketing and Consumer Research," *Journal of Consumer Research* (30:2), pp. 199-218.
- Jöreskog, K. G. 1983. "Factor Analysis as an Error-in-Variables Model," in *Principles of Modern Psychological Measurement: A Festschrift for Frederic M. Lord*, H. Wainer and S. Messick (eds.), Hillsdale, NJ: Lawrence Erlbaum Associates, pp. 185-196.
- Jöreskog, K. G., and Lawley, D. N. 1968. "New Methods in Maximum Likelihood Factor Analysis," *British Journal of Mathematical and Statistical Psychology* (21:1), pp. 85-96.
- Keil, M., Tan, B., Wei, K., Saarinen, T., Tuunainen, V., and Wassenaar, A. 2000. "A Cross-Cultural Study on Escalation of Commitment Behavior in Software Projects," *MIS Quarterly* (24:2), pp. 299-325.
- Kettenring, J. R. 1971. "Canonical Analysis of Several Sets of Variables," *Biometrika* (58:3), p. 433.
- Kline, R. B. 2010. *Principles and Practice of Structural Equation Modeling* (3rd ed.) New York: Guilford Press.
- Liang, H., Saraf, N., Hu, Q., and Xue, Y. 2007. "Assimilation of Enterprise Systems: The Effect of Institutional Pressures and the Mediating Role of Top Management," *MIS Quarterly* (31:1), pp. 59-87.

- Lohmöller, J.-B. 1989. *Latent Variable Path Modeling with Partial Least Squares*, Heidelberg: Physica.
- Lord, F., and Novick, M. 1968. *Statistical Theories of Mental Test Scores*, Reading, MA: Addison-Wesley.
- Lu, I. R. R., Kwan, E., Thomas, D. R., and Cedzynski, M. 2011. "Two New Methods for Estimating Structural Equation Models: An Illustration and a Comparison with Two Established Methods," *International Journal of Research in Marketing* (28:3), pp. 258-268.
- Marcoulides, G. A., Chin, W. W., and Saunders, C. 2012. "When Imprecise Statistical Statements Become Problematic: A Response to Goodhue, Lewis, and Thompson," *MIS Quarterly* (36:3), pp. 717-728.
- Marcoulides, G. A., and Saunders, C. 2006. "PLS: A Silver Bullet?," *MIS Quarterly* (30:2), pp. iii-ix.
- McDonald, R. P. 1996. "Path Analysis with Composite Variables," *Multivariate Behavioral Research* (31:2), pp. 239-270.
- Nunnally, J. C., and Bernstein, I. H. 1994. *Psychometric Theory* (3rd ed.) New York: McGraw-Hill.
- Paxton, P., Curran, P. J., Bollen, K. A., Kirby, J., and Chen, F. 2001. "Monte Carlo Experiments: Design and Implementation," *Structural Equation Modeling* (8:2), pp. 287-312.
- Qureshi, I., and Compeau, D. 2009. "Assessing Between-Group Differences in Information Systems Research: A Comparison of Covariance- and Component-Based SEM," *MIS Quarterly* (33:1), pp. 197-214.
- R Core Team. 2013. *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna (<http://www.R-project.org>).
- Rai, A., Goodhue, D., Henseler, J., and Thompson, R. 2013. "To PLS or Not to PLS: That Is the Question," in *Proceedings of the 19th Americas Conference on Information Systems*, Chicago, IL.
- Reinartz, W. J., Haenlein, M., and Henseler, J. 2009. "An Empirical Comparison of the Efficacy of Covariance-Based and Variance-Based SEM," *International Journal of Research in Marketing* (26:4), pp. 332-344.
- Rigdon, E. E. 1994. "Demonstrating the Effects of Unmodeled Random Measurement Error," *Structural Equation Modeling* (1:4), pp. 375-380.
- Rigdon, E. E. 2012. "Rethinking Partial Least Squares Path Modeling: In Praise of Simple Methods," *Long Range Planning* (45:5-6), pp. 341-358.
- Rigdon, E. E. 2014. "Comment on 'Improper Use of Endogenous Formative Variables,'" *Journal of Business Research* (67:1), pp. 2800-2802.
- Ringle, C. M., Sarstedt, M., and Straub, D. W. 2012. "Editor's Comments: A Critical Look at the Use of PLS-SEM in *MIS Quarterly*," *MIS Quarterly* (36:1), pp. iii-xiv.
- Rönkkö, M., and Ylitalo, J. 2010. "Construct Validity in Partial Least Squares Path Modeling," in *Proceedings of the 31st International Conference on Information Systems*, St. Louis, MO.
- Rosseel, Y. 2012. "lavaan: An R Package for Structural Equation Modeling," *Journal of Statistical Software* (48:2), pp. 1-36.
- Sarstedt, M., Ringle, C. M., and Hair, J. F. 2014. "PLS-SEM: Looking Back and Moving Forward," *Long Range Planning* (47:3), pp. 132-137.
- Sarstedt, M., Ringle, C. M., Henseler, J., and Hair, J. F. 2014. "On the Emancipation of PLS-SEM: A Commentary on Rigdon (2012)," *Long Range Planning* (47:3), pp. 154-160.
- Schönemann, P. H., and Steiger, J. H. 1976. "Regression Component Analysis," *British Journal of Mathematical and Statistical Psychology* (29:2), pp. 175-189.
- Shah, R., and Goldstein, S. M. 2006. "Use of Structural Equation Modeling in Operations Management Research: Looking Back and Forward," *Journal of Operations Management* (24:2), pp. 148-169.
- Sijtsma, K. 2009. "On the Use, the Misuse, and the Very Limited Usefulness of Cronbach's Alpha," *Psychometrika* (74:1), pp. 107-120.
- Tenenhaus, A., and Tenenhaus, M. 2011. "Regularized Generalized Canonical Correlation Analysis," *Psychometrika* (76:2), pp. 257-284.
- Tenenhaus, M. 2008. "Component-Based Structural Equation Modelling," *Total Quality Management & Business Excellence* (19:7), pp. 871-886.
- Vale, C. D., and Maurelli, V. A. 1983. "Simulating Multivariate Nonnormal Distributions," *Psychometrika* (48:3), pp. 465-471.
- Wetzels, M., Odekerken-Schröder, G., and van Oppen, C. 2009. "Using PLS Path Modeling for Assessing Hierarchical Construct Models: Guidelines and Empirical Illustration," *MIS Quarterly* (33:1), pp. 177-195.
- Wold, H. O. A. 1982. "Soft Modeling: The Basic Design and Some Extensions," in *Systems under indirect observation*, K. G. Jöreskog and H. O. A. Wold (eds.), Amsterdam: North-Holland, pp. 1-54.
- Yuan, K. H., and Bentler, P. M. 2002. "On Robustness of the Normal-Theory Based Asymptotic Distributions of Three Reliability Coefficient Estimates," *Psychometrika* (67:2), pp. 251-259.

About the Authors

Theo K. Dijkstra is a professor at the University of Groningen, The Netherlands. He was a Fulbright scholar at the University of California, Department of Psychology. He has published or has articles accepted in journals including *Information Systems Research*, *European Journal of Information Systems*, *Computational Statistics & Data Analysis*, *Econometrica*, *Journal of Econometrics*, *Organizational Research Methods*, *Psychometrika*, *The International Economic Review*, *British Journal of Mathematical and Statistical Psychology*, and others. He also worked for an institutional asset manager, responsible for the development of decision support tools and portfolio strategies. His research interests include statistical methodology, multivariate statistics, multicriteria decision analysis, and partial least squares.

Jörg Henseler is a professor at the University of Twente, The Netherlands, and a visiting professor at NOVA IMS, Universidade Nova de Lisboa, Portugal. He has published in scholarly journals including *European Journal of Information Systems*, *Computational Statistics & Data Analysis*, *International Journal of Research in Marketing*, *Journal of the Academy of Marketing Science*, *Organizational Research Methods*, and *Structural Equation Modeling*, among others, and he has edited two handbooks on partial least squares path modeling. His research interests encompass structural equation modeling, marketing research, and management of products, services, and brands.

CONSISTENT PARTIAL LEAST SQUARES PATH MODELING

Theo K. Dijkstra

Faculty of Economics and Business, University of Groningen, Nettelbosje 2,
9747 AE Groningen THE NETHERLANDS {t.k.dijkstra@rug.nl}

Jörg Henseler

Faculty of Engineering Technology, University of Twente, Drienerlolaan 5,
7522 NB Enschede THE NETHERLANDS {j.henseler@utwente.nl}
NOVA IMS, Universidade Nova de Lisboa, 1070-312 Lisbon PORTUGAL {jhenseler@isegi.unl.pt}

Appendix A

In variance-based SEM, the path coefficients can be determined based on the construct score correlation matrix \mathbf{R} , which for the model depicted in Figure 1 is of the following form (“cor” = correlation; “rel” = reliability):

$$\mathbf{R} = \begin{bmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \text{cor}(\xi_1, \xi_2) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\xi}_2)} & \text{cor}(\xi_1, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\eta})} \\ \text{cor}(\xi_1, \xi_2) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\xi}_2)} & 1 & \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_2) \cdot \text{rel}(\tilde{\eta})} \\ \text{cor}(\xi_1, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\eta})} & \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_2) \cdot \text{rel}(\tilde{\eta})} & 1 \end{bmatrix}$$

From \mathbf{R} , we can extract the submatrix $\mathbf{R}_x = \begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix}$

and the subvector $\mathbf{R}_{xy} = \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix}$.

Under the assumption that the construct scores are standardized, the regression equation equals

$$\hat{\beta} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = \mathbf{R}_x^{-1} \mathbf{R}_{xy} = \begin{bmatrix} 1 & r_{12} \\ r_{12} & 1 \end{bmatrix}^{-1} \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} = \frac{1}{\det(\mathbf{R}_x)} \begin{bmatrix} 1 & -r_{12} \\ -r_{12} & 1 \end{bmatrix} \begin{bmatrix} r_{13} \\ r_{23} \end{bmatrix} = \frac{1}{1-r_{12}^2} \begin{bmatrix} r_{13} - r_{12}r_{23} \\ r_{23} - r_{12}r_{13} \end{bmatrix}$$

Since the true β_1 equals zero, we can exploit that $\text{cor}(\xi_1, \eta) = \text{cor}(\xi_1, \xi_2) \cdot \text{cor}(\xi_2, \eta)$ and thus find the following result for the path $\hat{\beta}_1$ coefficient as estimated by variance-based SEM:

$$\begin{aligned}
\hat{\beta}_1 &= \frac{1}{1 - r_{12}^2} \cdot (r_{13} - r_{12}r_{23}) \\
&= \frac{\text{cor}(\xi_1, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\eta})} - \text{cor}(\xi_1, \xi_2) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\xi}_2)} \cdot \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_2) \cdot \text{rel}(\tilde{\eta})}}{1 - \text{cor}^2(\tilde{\xi}_1, \tilde{\xi}_2)} \\
&= \frac{\text{cor}(\xi_1, \xi_2) \cdot \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\eta})} - \text{cor}(\xi_1, \xi_2) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\xi}_2)} \cdot \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_2) \cdot \text{rel}(\tilde{\eta})}}{1 - \text{cor}^2(\tilde{\xi}_1, \tilde{\xi}_2)} \\
&= \frac{\text{cor}(\xi_1, \xi_2) \cdot \text{cor}(\xi_2, \eta) \cdot \sqrt{\text{rel}(\tilde{\xi}_1) \cdot \text{rel}(\tilde{\eta})} \cdot (1 - \text{rel}(\tilde{\xi}_2))}{1 - \text{cor}^2(\tilde{\xi}_1, \tilde{\xi}_2)} \\
&= \frac{\text{cor}(\tilde{\xi}_1, \tilde{\xi}_2)}{1 - \text{cor}^2(\tilde{\xi}_1, \tilde{\xi}_2)} \cdot \text{cor}(\tilde{\xi}_2, \tilde{\eta}) \cdot \frac{1 - \text{rel}(\tilde{\xi}_2)}{\text{rel}(\tilde{\xi}_2)}
\end{aligned}$$

Appendix B¹

A starting point for PLS-analyses is the so-called basic design, in essence a factor model. It is assumed that we have N i.i.d.² column vectors of observed scores $y_1, y_2, y_3, \dots, y_N$ (i.e., we have N observations). We assume the usual standardization in PLS of each indicator, which means that each vector y_n for $n = 1, 2, 3, \dots, N$ has zero mean and its elements have unit variance. The vectors can be partitioned into M subvectors, $M \geq 2$, each with at least two components. For the i^{th} subvector y_{in} of y_n , we have

$$y_{in} = \lambda_i \cdot \eta_{in} + \varepsilon_{in} \quad (1)$$

where the loading vector λ_i and the vector of measurement errors ε_{in} have the same dimensions as y_{in} , and the unobservable latent variable η_{in} is real-valued.³ For convenience, we make the sufficient but by no means necessary assumption that all components of all error vectors are mutually independent, and independent of all latent variables.

The latent variables have also zero mean and unit variance. The correlation between η_{in} and η_{jn} will be denoted by ρ_{ij} . A particular set of easy implications is that the covariance matrix Σ_{ii} of y_{in} can be written as

$$\Sigma_{ij} := E y_{ij} y_{jn}^T = \rho_{ij} \lambda_i \lambda_j^T \quad (2)$$

where the covariance matrix of the measurement errors Θ_i is diagonal with non-negative diagonal elements, and we have for the covariance between y_{in} and y_{jn}

¹To keep the paper self-contained, we collect here the main algebraic results underlying PLSc (for a more elaborate explication, see Dijkstra 2010). Note that there are many ways to correct for inconsistency and the one we use here is quite possibly the simplest one. See Dijkstra (2013) for alternatives and extensions. Any combination of PLS with one of the alternatives could legitimately be called PLSc.

²The use of i.i.d. (independent and identically distributed) signifies that the data are seen as a random sample from a population. This is for conceptual convenience only. Together with the classical laws of large numbers and the central limit theorem (Cramér 1946), this entails the consistency and asymptotic normality properties we need. We can make much more general assumptions, but the statistical subtleties would be a distraction and would not lead to different results (CAN-estimators).

³We follow exactly the specification of Joreskog (1969). Please note that variables are arranged per rows, and observations per columns. Moreover, note that, as opposed to our custom in the main body of the paper, it pays here to use a subscript for the blocks.

$$\Sigma_{ij} := E y_{in} y_{jn}^T = \rho_{ij} \lambda_i \lambda_j^T \quad (3)$$

The covariance matrix of each y_n will be denoted by Σ , and the sample covariance matrix by S . The sample counterparts of Σ_{ii} and Σ_{ij} are denoted by S_{ii} and S_{ij} , respectively. Since we assume that the sample data are standardized before being analyzed, we have $S_{ij} = \frac{1}{N} \sum_{n=1}^N y_{in} y_{jn}^T$. Note that the assumptions made so far entail that the sample counterparts are consistent and jointly asymptotically normal estimators of the theoretical variance and covariance matrices (Cramér 1946, Chapter 28).

PLS features a number of iterative *fixed-point* algorithms, of which we select one, the so-called Mode A algorithm. This is, in general, numerically the most stable algorithm.⁴ As a rule, it converges from arbitrary starting vectors, and it is usually very fast. The outcome is an estimated weight vector \hat{w} , with typical subvector \hat{w}_i of the same dimensions as y_{in} . With these weights, *sample proxies* are defined for the latent variables:

$\hat{\eta}_{in} := \hat{w}_i^T y_{in}$ for η_{in} , with the customary normalization of a unit sampling variance, so $\frac{1}{N} \sum_{n=1}^N (\hat{w}_i^T y_{in})^2 = \hat{w}_i^T S_{ii} \hat{w}_i = 1$. In

Wold's PLS approach the $\hat{\eta}_{in}$'s replace the unobserved latent variables, and loadings and structural parameters are estimated using ordinary least squares.⁵ For Mode A, we have for each i (\propto stands for "proportional to")

$$\hat{w}_i \propto \sum_{j \in C(i)} \text{sign}_{ij} \cdot S_{ij} \hat{w}_j \quad (4)$$

Here, sign_{ij} is the sign of the sample correlation between $\hat{\eta}_i$ and $\hat{\eta}_j$, and $C(i)$ is a set of indices of latent variables. Traditionally, $C(i)$ contains the indices of latent variables adjacent to $\hat{\eta}_i$ (i.e., the indices of latent variables that appear on the other side of the structural or path equations in which $\hat{\eta}_i$ appears). Clearly, \hat{w}_i is obtained by a regression of the indicators y_{in} on the sign-weighted sum: $\sum_{j \in C(i)} \text{sign}_{ij} \cdot \hat{\eta}_{jn}$. There are other versions (with correlation weights, for example); this is one of the very simplest, and it is the original one (see Wold, 1982). There is little motivation in the PLS literature for the coefficients of $S_{ij} \hat{w}_j$,⁶ but the particular choice can be shown to be irrelevant for the probability limits of the estimators. The algorithm takes an arbitrary starting vector, and then basically follows the sequence of regressions for each i , each time inserting updates when available (or after each full round; the precise implementation is not important).

Dijkstra (1981, 2010) has shown that the PLS modes converge with a probability tending to one when the sample size tends to infinity, for essentially arbitrary starting vectors. Moreover, the weight vectors that satisfy the fixed-point equations, are locally continuously differentiable functions of the sample covariance matrix of y . They, as well as other estimators that depend smoothly on the weight vectors and S , are therefore jointly asymptotically normal.

Let us denote the probability limit of \hat{w}_i , $\text{plim } \hat{w}_i$, by \bar{w}_i . We can get it from the equation for \hat{w}_i by substitution of Σ for S .

$$\bar{w}_i \propto \sum_{j \in C(i)} \text{sign}_{ij} \cdot \sum_j \bar{w}_j = \sum_{j \in C(i)} \text{sign}_{ij} \cdot \rho_{ij} \lambda_i \lambda_j^T \bar{w}_j = \lambda_i \sum_{j \in C(i)} \text{sign}_{ij} \cdot \rho_{ij} \lambda_j^T \bar{w}_j \quad (5)$$

Since $\lambda_i^T \bar{w}_j$ is a scalar, and all terms in the sum have the vector λ_i in common, it is clear that $\bar{w}_i \propto \lambda_i$. We must have $\bar{w}_i^T \sum_{ii} \bar{w}_i = 1$, so

$$\bar{w}_i = \frac{\lambda_i}{\sqrt{\lambda_i^T \sum_{ii} \lambda_i}} \quad (6)$$

⁴This is because Mode A ignores collinearity between the observed variable predictors of the proxy, while Mode B takes account of that collinearity and thus must find the inverse of the covariance matrix of the predictions, which itself is sometimes unstable.

⁵Other estimators like 2SLS are also possible.

⁶In favor of the sign-weights, one could argue that in sufficiently large samples $\text{sign}_{ij} \cdot S_{ij} \hat{w}_j$ is approximately equal to $\lambda_j \cdot |\rho_{ij}| \cdot (\lambda_j^T \bar{w}_j)$, where the term in brackets measures the (positive) correlation between η_j and its proxy; see below for results that help justify this claim. So the tighter the connection between η_j and the better η_j can be measured, the more important $\hat{\eta}_j$ is in determining \hat{w}_i .

We conclude that PLS, Mode A, produces estimated weight vectors that tend to vectors proportional to the true loadings. One would like to have a simple estimate for the proportionality factor. We propose here as in Dijkstra (1981, 2010, 2011) to define $\hat{\lambda}_i := \hat{c}_i \cdot \hat{w}_i$, where the scalar \hat{c}_i is such that the *off-diagonal* elements of S_{ii} are reproduced as well as possible in a least squares sense. So we minimize the Euclidean distance between⁷

$$[S_{ii} - \text{diag}(S_{ii})] \text{ and } [(c_i \cdot \hat{w}_i)(c_i \cdot \hat{w}_i)^T - \text{diag}((c_i \cdot \hat{w}_i)(c_i \cdot \hat{w}_i)^T)] \quad (7)$$

as a function of c_i and obtain

$$\hat{c}_i := \left[\frac{\hat{w}_i^T (S_{ii} - \text{diag}(S_{ii})) \hat{w}_i}{\hat{w}_i^T (\hat{w}_i \hat{w}_i^T - \text{diag}(\hat{w}_i \hat{w}_i^T)) \hat{w}_i} \right]^{\frac{1}{2}} \quad (8)$$

The use of “diag” means that only the off-diagonal elements of the matrices are taken into account. So the numerator within brackets is just $\sum_{a \neq b} \hat{w}_{ia} \hat{w}_{ib} S_{ii,ab}$ and similarly for the denominator. In sufficiently large samples, \hat{c}_i will be well-defined, real, and positive. (In all samples in this paper and those in other studies, \hat{c}_i attained proper values.) Its calculation does not require additional numerical optimization. It is straightforward to verify, by replacing S_{ii} by Σ_{ii} and \hat{w}_i by \bar{w}_i , that the correction does its job: the matrix in the denominator equals the matrix in the numerator, apart from a factor $\frac{1}{\lambda_i^T \sum_{ii} \lambda_i}$, so

$$\bar{c}_i := \text{plim } \hat{c}_i = \sqrt{\lambda_i^T \sum_{ii} \lambda_i} \quad (9)$$

Now, in particular

$$\text{plim } \hat{\lambda}_i = \text{plim } (\hat{c}_i \cdot \hat{w}_i) = \bar{c}_i \cdot \bar{w}_i = \lambda_i \quad (10)$$

It will be useful to define a population proxy $\bar{\eta}_{in}$ by $\bar{\eta}_{in} := \bar{w}_i^T y_{in}$. Clearly, the squared correlation between a population proxy and its corresponding latent variable is

$$R^2(\eta_i, \bar{\eta}_i) = (\bar{w}_i^T \lambda_i)^2 \quad (11)$$

which equals

$$(\lambda_i^T \lambda_i)^2 \div \lambda_i^T \Sigma_{ii} \lambda_i = \frac{(\lambda_i^T \lambda_i)^2}{(\lambda_i^T \lambda_i)^2 + \lambda_i^T \Theta_i \lambda_i} \quad (12)$$

With a large number of high quality indicators, this correlation can be close to one (“consistency at large” in PLS parlance). A trivially deduced but important algebraic relationship is

$$R^2(\bar{\eta}_i, \bar{\eta}_j) = (\bar{w}_i^T \Sigma_{ij} \bar{w}_j)^2 = \rho_{ij}^2 \cdot R^2(\eta_i, \bar{\eta}_i) \cdot R^2(\eta_j, \bar{\eta}_j) \quad (13)$$

indicating that the PLS proxies will tend to underestimate the squared correlations between the latent variables. In fact, one can show that this is true for multiple correlations as well (see Dijkstra 2010). Also note that

$$R^2(\eta_i, \bar{\eta}_j) = (\bar{w}_i^T \lambda_j)^2 = (\bar{w}_i^T \cdot (\bar{w}_i \cdot \bar{c}_i))^2 = (\bar{w}_i^T \bar{w}_i)^2 \cdot \bar{c}_i^2 \quad (14)$$

so that we can estimate the (squared) quality of the proxies consistently by

⁷This assumes that the measurement errors within a block are uncorrelated, the basic design. If we know that some errors are correlated or we have doubts about them, we can delete the items from the difference to be minimized.

$$\hat{R}^2(\eta_i, \bar{\eta}_i) := (\hat{w}_i^T \hat{w}_i)^2 \cdot \hat{c}_i^2 \quad (15)$$

Moreover, with

$$\hat{R}^2(\bar{\eta}_i, \bar{\eta}_j) := (\hat{w}_i^T S_{ij} \hat{w}_j)^2 \cdot \hat{c}_i^2 \quad (16)$$

we can estimate the correlations between the latent variables consistently (see Equation 13)

$$\hat{\rho}_{ij} := \sqrt{\frac{\hat{R}^2(\bar{\eta}_i, \bar{\eta}_j)}{\hat{R}^2(\eta_i, \bar{\eta}_i) \cdot \hat{R}^2(\eta_j, \bar{\eta}_j)}} \quad (17)$$

We close this appendix with four observations:

1. Standard PLS software for Mode A will produce all the necessary, simple ingredients for consistent estimation.
2. The approach can be and has been extended to Mode B (Dijkstra 1981, 2010, 2011), but since Mode A is numerically more stable and faster than Mode B, it has first priority.
3. In the main body of this paper, we used $\rho(\eta_i)$ for $R(\eta_i, \bar{\eta}_i)$.
4. In the main body of this paper, we used (without subscript i)

$$\rho_{A,i} := (\hat{w}_i^T \hat{w}_i)^2 \cdot \frac{\hat{w}_i^T (S_{ii} - \text{diag}(S_{ii})) \hat{w}_i}{\hat{w}_i^T (\hat{w}_i, \hat{w}_i^T - \text{diag}(\hat{w}_i, \hat{w}_i^T)) \hat{w}_i} \quad (18)$$

This is just $\hat{R}^2(\eta_i, \bar{\eta}_i)$.

References

- Cramér, H. 1946. *Mathematical Models of Statistics* (Volume 9), Princeton, NJ: Princeton University Press.
- Dijkstra, T. K. 1981. *Latent Variables in Linear Stochastic Models: Reflections on "Maximum Likelihood" and "Partial Least Squares" Methods*, Ph.D. Thesis, Groningen University. (A second edition was published in 1985 by Sociometric Research Foundation, Amsterdam.)
- Dijkstra, T. K. 2010. "Latent Variables and Indices: Herman Wold's Basic Design and Partial Least Squares," in *Handbook of Partial Least Squares: Concepts, Methods, and Applications*, V. E. Vinzi, W. W. Chin, J. Henseler, and H. Wang (eds.), New York: Springer, pp. 23-46.
- Dijkstra, T. K. 2011. "Consistent Partial Least Squares Estimators for Linear and Polynomial Factor Models," working paper, University of Groningen, Groningen, The Netherlands (<http://www.rug.nl/staff/t.k.dijkstra/research>).
- Dijkstra, T. K. 2013. "A Note on How to Make PLS Consistent," working paper, University of Groningen, Groningen, The Netherlands (<http://www.rug.nl/staff/t.k.dijkstra/how-to-make-pls-consistent.pdf>).
- Jöreskog, K. G. 1969. "A General Approach to Confirmatory Maximum Likelihood Factor Analysis," *Psychometrika* (34:2), pp. 183-202.
- Wold, H. O. A. 1982. "Soft Modeling: The Basic Design and Some Extensions," in *Systems under indirect observation*, K. G. Jöreskog and H. O. A. Wold (eds.), Amsterdam: North-Holland, pp. 1-54.